Technical Report 1061

Assessment of User Reactions to the Multi-Service Distributed Training Testbed (MDT2) System

Angelo Mirabella
U.S. Army Research Institute

Paul Sticha
Human Resources Research Organization

John Morrison
Institute for Defense Analysis

April 1997

United States Army Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Technical review by

Richard E. Christ
J. Douglas Dressel

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED (from. . . to)
1997, April	Final	July 1993-June 1996

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
Assessment of User Reactions to the Multi-Service Distributed Training Testbed (MDT2) System	5b. PROGRAM ELEMENT NUMBER 0603007A
6. AUTHOR(S)	5c. PROJECT NUMBER A793
Angelo Mirabella (ARI), Paul Sticha (HumRRO), and John Morrison (IDA)	5d. TASK NUMBER 2122
	5e. WORK UNIT NUMBER H01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences ATTN: PERI-II 5001 Eisenhower Avenue Alexandria, VA 22333-5600	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600	10. MONITOR ACRONYM ARI
	11. MONITOR REPORT NUMBER Technical Report 1061

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT *(Maximum 200 words)*:

This research was part of a larger program to develop a methodology for multi-Service training of Close Air Support (CAS), using Distributed Interactive Simulation (DIS) technology. This paper summarizes an assessment of user reactions to the training. Survey questionnaires, group interviews, and observations of the training were used to obtain data on two questions. What value is added to existing Service training cycles by the DIS methodology developed? How well did the training work? A key finding was that the distributed methodology fills a critical gap in training multi-Service CAS coordination tasks. An important potential application of the research is to "ramp-up" training in preparation for rotations to combat training centers. Many lessons were learned about how to develop and apply survey and interview instruments as part of a larger evaluation of DIS training. For example, we "discovered" that for multiple sites and services "one size does not fit all." Assessment instruments must be carefully prepared in different versions to suit the varying perspectives and roles of multiple services.

15. SUBJECT TERMS

Unit training	Performance assessment	Collective training	Distributed Interactive Simulation
Networked simulation	Distance learning	Training program evaluation	Program evaluation

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT	17. ABSTRACT	18. THIS PAGE			
Unclassified	Unclassified	Unclassified	Unlimited	105	Dr. Angelo Mirabella (703) 617-0325

Technical Report 1061

Assessment of User Reactions to the Multi-Service Distributed Training Testbed (MDT2) System

Angelo Mirabella
U.S. Army Research Institute

Paul Sticha
Human Resources Research Organization

John Morrison
Institute for Defense Analysis

Advanced Training Methods Research Unit
Robert J. Seidel, Chief

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

April 1997

Army Project Number
2O363007A793

Training Systems and Education

FOREWORD

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) conducts research on how to design unit training strategies. Within the past few years, its mission has been extended to include the Army's role in multi-Service training. This document is one of a number of reports that have been produced under the Multi-Service Distributed Training Testbed (MDT2) program. The purpose of the program was to develop and try out methods for planning and executing inter-Service distributed interactive simulation-based training of close air support at the battalion task force level.

The present report summarizes the results of an assessment of user reactions to the MDT2 training system. The assessment was part of a broader effort to estimate the value added by MDT2 and to examine the usefulness of a number of training and feedback tools. Overall, the assessment showed that troops perceived the methodology to provide valued training in inter-Service coordination tasks.

Portions of the material in this report have been included as recommendations in a report to the Defense Modeling and Simulation Office (DMSO) sponsor and funding agency for the program.

ZITA M. SIMUTIS
Technical Director

EDGAR M. JOHNSON
Director

ACKNOWLEDGMENTS

The MDT2 program was a consummate example of many people working together and helping one another to support the Services and national security. In particular, the assessment portion of the program - the focus of this report - would not have been possible without the invaluable contributions of many people during planning and conduct of the effort. Our gratitude extends to Dr. Maize Knerr, Human Resources Research Organization, (HumRRO); Dr. Frank Moses, Army Research Institute (ARI); Dr. Herbert Bell and Becky Brooks, Armstrong Laboratory (AL); Randy Oser, Dr. Dan Dwyer, Dr. Ruth Willis, and Dr. Eduardo Salas, Naval Air Warfare Center Training Systems Division (NAWTCTSD). Special thanks are owed to Randy, Becky, and Ruth for collecting survey, interview, and activity data at the disparate sites comprising the MDT2 network. Their generous efforts made it possible to accomplish the mission, with limited resources.

We're indebted also to the peer reviewers, Douglas Dressel and Dr. Richard Christ, who read the manuscript with meticulous and pains-taking care. They made many suggestions for improving its readability, and technical and editorial accuracy.

ASSESSMENT OF USER REACTIONS TO THE MULTI-SERVICE DISTRIBUTED TRAINING TESTBED (MDT2) SYSTEM

EXECUTIVE SUMMARY

Research Requirement:

Assess the value added to existing Service training of Close Air Support (CAS) Multi-Service Distributed Interactive Training; Determine ways to improve the planning, management, and conduct of such training; and, develop methodology for conducting surveys and interviews as part of larger evaluations of distributed training.

Procedure:

Survey questionnaires, group interviews, and observations of the training were used to obtain data on two questions. What value is added to existing Service training cycles by the Distributed Interactive Simulation DIS methodology developed? How well did the training work?

Findings:

A key finding was that the distributed methodology fills a critical gap in training multi-Service CAS coordination tasks. An important potential application of the research is to 'ramp-up' training in preparation for rotations to Combat Training Centers (CTCs). Many lessons were learned about how to develop and apply survey and interview instruments as part of a larger evaluation of distributed interactive simulation training. For example, we 'discovered' that for multiple sites and services 'one size does not fit all'. Assessment instruments must be carefully prepared in different versions to suit the varying perspectives and roles of multiple services.

Utilization of Findings:

This report can serve as a source document in the development and test of prototype DIS training systems. Furthermore lessons learned about how to assess user reactions will contribute to the Army's general technology base on training system evaluation. The report's findings about the value of DIS for multi-Service training can assist Army and inter-service policy making concerning the acquisition and use of DIS-based training systems.

ASSESSMENT OF USER REACTIONS TO THE MULTI-SERVICE DISTRIBUTED TRAINING (MDT2) TRAINING SYSTEM

CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Assessment of User Reactions to the Multi-Service Distributed Training Testbed (MDT2) System

INTRODUCTION

This report documents an assessment of the Multi-Service Distributed Training Testbed Program (MDT2-P). The report deals with the value added by the program's methodology to planning and conducting Multi-Service training of Close Air Support (CAS). The Testbed included armored vehicle, F16, forward air control, and laser targeting simulators which were geographically distributed but electronically linked. We used it to develop and evaluate instructional principles and tools for distributed tactical training. An important source of data for this development was provided by assessment of opinions and ratings of participants in the MDT2 training exercises. This report documents that assessment.

The report details the questions asked, the data collection instruments and procedures, and findings of an assessment of user reactions to the training. It also addresses ways to improve assessment methodology and to use that methodology to support further developments of multi-Service Distributed Interactive Simulation (DIS) training. Portions of the assessment results are documented in Mirabella (1995). Other aspects of the MDT2 program are documented in Bell (1995); Dwyer, Oser, Fowlkes, and Meliza (1995), Hawley and Christ (in press), and Moses (1995). A comprehensive description of the program, including practical recommendations on training system assessment appears in a Four-Service Project Technical Report (Department of Defense, 1996). A cost-effectiveness assessment is documented in Orlansky, Taylor, and Levine (1996).

The assessment of user reactions complemented checklist measures of unit performance and CAS mission outcomes (Dwyer, Oser, Fowlkes, & Meliza, 1995). Each type of measure is needed for a total training system evaluation (Chen, 1995; Crego, 1994; Pinker, Samuel, & Batcher, 1995). We need performance data to answer several questions. Did expected coordinations among the Service elements occur? How well were they carried out? Did these measures improve with repeated exercises? If not, why not? Equally important were the personal reactions of the troops and observer/controllers (O/Cs)in the training exercises.

Data on user reactions to new training systems are important. Only satisfied customers will use simulation to train no matter how effective it is in improving or sustaining performance. The customer also has valuable insights into needed product improvement.

An additional reason for data on user reactions merits special attention. User reaction data are a critical precursor to measurement of task proficiency and cost-effectiveness analysis (CTEA). Without such data to help 'debug' training systems, CTEA could be a "re-arrangement of deck chairs on the Titanic" particularly during early development of a training system when serious deficiencies in system design or application otherwise may be overlooked.

The distributed, multi-Service environment requires even greater 'early vigilance' than single site training. Some of the users and system components will be physically separated from the principal training developers and managers. Overlooked problems at one site can (and did in the second year of this program)[1] have severe repercussions at other sites. [Note, superscripted numbers refer to End Notes appearing after the appendices.]

MDT2 TRAINING TESTBED ENVIRONMENT

MISSIONS

Armored tank battalion attack and defend operations with CAS were used as testbed training missions. Close Air Support was supported by a Marine laser detection unit which acquired enemy targets and guided ordnance from F16s to those targets. Each attack mission began with a movement to contact by a live company team and two simulated companies. Part way through the mission, a pair of F16s provided CAS. Each defensive mission started with pre-planned, prepared positions.

MDT2 Components

The testbed included the following networked, distributed facilities (Figure 1). Each location used a comparable terrain data base which simulated an area at the National Training Center (NTC), Ft. Irwin, California.

Simulators for an Armored Battalion Task Force 'slice' at the Mounted Warfare Testbed, Fort Knox, KY. A tactical operations center (TOC) included an Operations Officer (S3), an Intelligence Officer (S2). It also contained a Fire Support Element-Non Commissioned Officer (FSE-NCO) and an Air Force Tactical Control Party (TACP). The Battalion (Bn) Commander, Air Liaison Officer (ALO), and Fire Support Officer (FSO) occupied armored vehicle simulators. A Company Team Commander, Executive Officer, Fire Support Team (FIST), and Scout occupied the remaining simulated vehicles with drivers and gunners. Additional vehicles needed to complete the 'live' company were simulated.

In the first of two developmental tests in May 1994, the Army participants were from the Kentucky National Guard. In the second test, February 1995, the Army participants were from an active brigade at Fort Hood.[2] Enemy ground forces, their actions, and reactions were computer-generated using the Modified Semi-Automated Forces (MODSAF) program.

F-16 Block 30 aircraft simulators for attack pilots at Armstrong Laboratory, Mesa, AZ. Four Reserve Air Force Pilots served as players. Three of these participated in the 1994 and the 1995 sets of exercises. They also took part in Air Warrior out of Nellis AFB between the MDT2 sets. A pair of aircraft were employed for each mission.

Laser designator simulators, operated by the Marine Corps at San Diego, CA in 1994[3] and Armstrong Laboratory in 1995[4]. These simulated the Deployable Forward Observer/Modular Universal Laser Equipment (DFO/MULE). DFO/MULE was used to acquire targets for the F16s.

OV 10/22 cockpit simulator in 1994; Helmet-mounted display (HMD) simulating an OV10/22s for the Airborne Forward Air Controller (AFAC) in 1995. This was located at the Navy's Aircrew Systems Department, Naval Air Warfare Center, Patuxent River, MD. The players were Marine Corps pilots.

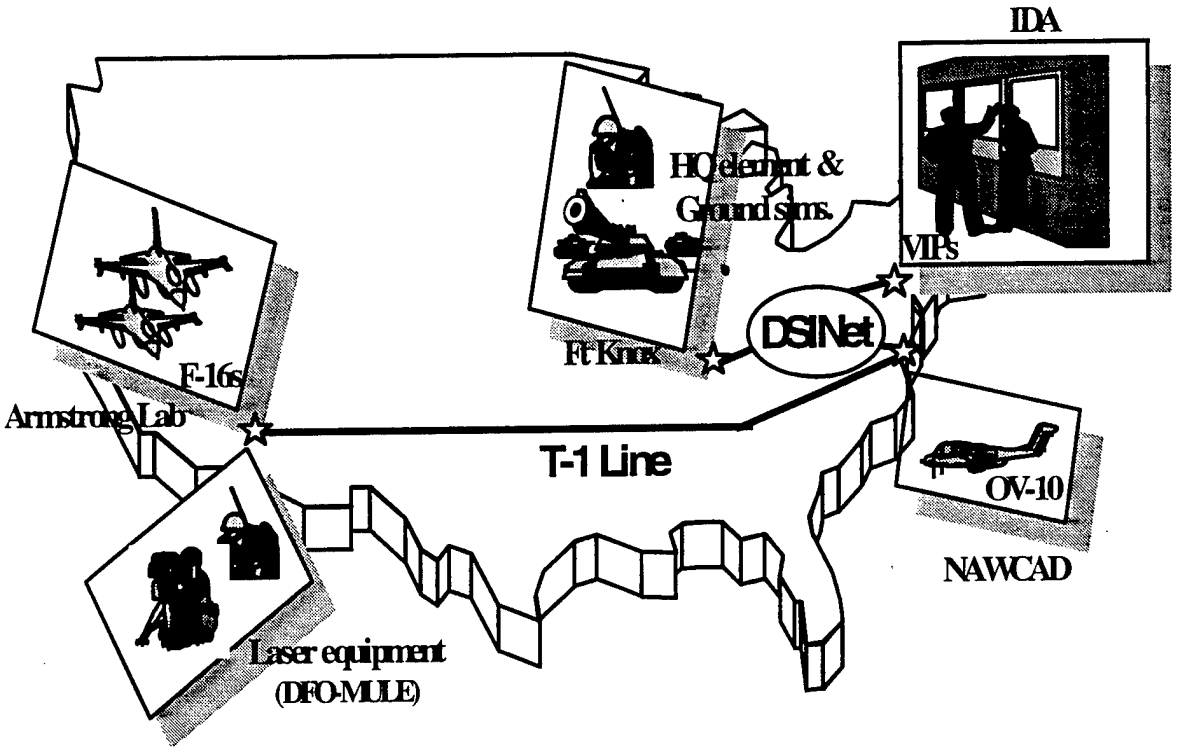

Figure 1. Components Of The MDT2 Testbed With an Additional Network Site for Visitor Observation at the Institute For Defense Analysis.

RESEARCH AND DEVELOPMENT APPROACH AND BACKGROUND

Approach

The MDT2 program employed case study with a sample size of two. Each of the sample elements was a week of simulation-based training exercises (SIMEXs), one in May, 1994, the other in February, 1995. Each daily SIMEX was followed by the preparation and delivery an after action review (AAR). Simulators at the various sites (Figure 1) used the same NTC terrain site, tailored (i.e., scaled) for ground and air operations. A variety of Research and Development (R & D) and support activities by the three Service laboratories (ARI, NAWCTSD, and AF/AL) ran concurrently with training activities.

Case study was appropriate for the mission of MDT2-P: develop, test, and deliver an experimental DIS-based methodology for training multi-Service combat tasks. The MDT2-P provided a rich opportunity for identifying researchable issues. But controlled experimentation was neither feasible, nor appropriate. Instead, project staff evolved candidate training methods and strategies, based on their training research and engineering expertise. In any event, a philosophical case for case study is suggested by the following:

> "So far as the laws of mathematics refer to reality, they are not certain. And so far as they are certain, they do not refer to reality." (Albert Einstein, Geometry and Experience; cited in Kosko, 1993, p. 3).

> "You find good math models only in textbooks and classrooms. They are toy answers to toy problems. The real world pays no attention to them." (Kosko, 1993, p. 169)

Research Background in Assessment of Training System Effectiveness

The training system assessment literature has addressed a large set of assessment measures. These include trainee satisfaction, skill acquisition, transfer of training, changes in job behavior and attitudes, and increases in 'productivity' (Hamblin, 1974; Kirkpatrick, 1967, 1976, 1987; Tannenbaum, Cannon-Bowers, Salas, & Mathieu, 1993). Pioneering work was done by Kirkpatrick and by Hamblin.

In his original, seminal research Kirkpatrick (1967) defined four variables and related data collection instruments: trainee reactions, learning, job behavior, and organizational results. He reiterated these variables and instruments in later publications with some revision of the instruments and updated illustrations of their use (Kirkpatrick, 1976, 1987).

Hamblin modified Kirkpatrick's approach to training system assessment by recommending that objectives (expected effects) of the training system be defined and actual effects measured for each variable. Hamblin's work is the basis for what is now referred to as 'Goal-Oriented' evaluation.

Knott (1994, pp. 230 - 239) cautioned against the goal-oriented approach, particularly for distributed training program development. Her comments are germane for the MDT2 Program. A focus on specific goals may result in overlooking unanticipated effects. Furthermore, goals may reflect the needs of some but not other stakeholders and customers. She recommends a practical compromise. Cautiously state some goals to satisfy sponsors, but 'spread a wide net' for unanticipated effects.

More recent research, sponsored by NAWCTSD, has examined additional variables, interactions among those variables, and measurement instruments beyond those provided by Kirkpatrick (Cannon-Bowers, Salas, Tannenbaum, & Mathieu, 1995; Kraiger, Ford, & Salas, 1993; Tannenbaum, Cannon-Bowers, Salas, & Mathieu, 1993).

The NAWCTSD research has made two major contributions to training assessment theory and practice. It has surfaced non-traditional assessment variables including trainee expectations, self-efficacy, motivation, cognitive ability, and organizational commitment. It has also modeled and validated interactions and sequential effects among measures of training program value. For example, an effective training program fosters 'can-do' attitude. This in turn improves performance (Tannenbaum et al., 1993).

The foregoing research on training program evaluation provided a point of departure for designing the MDT2 training system assessment. But, we needed to adapt the Kirkpatrick and more recent approaches to assessment. Kirkpatrick's approach was developed for use with content-oriented, individual, classroom, lecture-oriented training for industrial applications. His guidelines and assessment instruments are most suitable for that environment.

In contrast, MDT2 training was task oriented, collective, distributed, and simulation-based. The recent extensions of Kirk-patrick's work by NAWCTSD were too complicated and their data collection requirements too extensive for MDT2-P. But we borrowed key ideas from them.

Major Questions

Question 1. What estimated value can the MDT2 training system (MDT2-TS) add to the military training pipeline? Training value depends upon a number of subordinate dimensions. These were used to guide the design of data collection and analysis. The dimensions are:

a. Need for this type of training

b. Credibility and realism

c. Multi-Service value

d. Role in the training pipeline

e. Expected impact

Question 2. How well does the testbed system technology work for multi-Service training of close air support? We focused on two subsidiary questions to derive measures for Question 2.

a. How well were close air support training objectives (TOs) covered by MDT2 exercises? Twenty-five (25) TOs were compiled for use in designing the testbed exercises.

b. How useful were the instructional tools and methods that were selected or developed to support MDT-2 training? These tools and methods dealt with training development and management, and after action reviews (See Table 14 for a list of issues).

A similar but alternative framework for evaluation was evolved by Hawley and Christ (in press) from their observations and analyses of the MDT2 (1995) exercises. They begin with the following general definition:

> ***"A training system has value added vis-a-vis another if it (1) is capable of training skills that it is not possible to train using conventional methods or (2) ... provides a higher absolute level of performance than alternatives are capable of producing or results in equivalent performance at a lower cost ... "***

They add that CEA at least narrowly and technically defined, is appropriate to Part (2) of their general definition. In that case each alternative training system is capable of training the requisite skills, but at different costs. Hawley and Christ further qualify their definition with the following:

Another aspect of the value-added issue, and one that is in some sense prerequisite for quantifying value added, concerns the conditions under which a high-technology option is a suitable solution for a given performance and training situation.

Value-added, is relative to a very specific purpose and set of conditions. This appears to be a critical qualification. It suggests that careful attention needs to be paid to how the new training is used and for what specific purpose. Boldovici and Kolasinski (1996) have expressed a similar view in describing statistical methods for assessing device-based training effectiveness.

Hawley and Christ operationalize value-added by asking four questions: 1. Is the potential application ***suitable***? 2. Is it ***effective***? 3. Is it ***superior*** to conventional training methods? 4. Is the superiority operationally ***relevant***? They detail and provide measurement concepts for each of the questions.

In their test of the model, they address two assessment dimensions: suitability and estimated effectiveness They conclude that MDT2-TS is suitable for its intended purpose if it is not used for introductory CAS training, or as a complete substitute for field training. They also conclude that MDT2-TS is potentially effective in training multi-Service CAS synchronization and coordination tasks. We will return to their results in our discussion for Question 1.

METHODOLOGY

Overview of Research Design

May, 1994 exercises. CAS exercises were conducted during two 1-week sessions, separated by a week. The first session was a rehearsal. The intervening week was a time for fixing problems. The third week was the MDT2 demonstration/test. Early in Week 1, after participants had become oriented to the MDT2 environment, we administered site-tailored biographical surveys (Appendix A). These, and a Training Value Questionnaire, provided baselines on participant background and initial impressions of training value (Question 1). The value questionnaire was administered again at the end of Week 3 (Simulation Exercise [SIMEX] Week) along with additional questions on estimated training effectiveness. Thirty-one (31) observer/controllers (O/Cs) and players filled out surveys (9 O/Cs, 22 players).

At the end of SIMEX Week, we administered a Training Objectives Survey (Appendix A). This instrument was designed to evaluate how well training objectives had been covered by the MDT2 exercises (Question 2A). We also interviewed players and O/Cs in small groups. Open-ended questions were designed to elicit opinions that would supplement the survey data on training value and effectiveness (Question 1). Additional questions addressed the usefulness and effectiveness of training tools and methods (Question 2B). Interviewers at each site were provided with general, written instructions for conducting the interview. They were also given a list of questions to ask.

February, 1995 exercises. We replicated the 1994 research design with changes explained below. Thirty-one (31) O/Cs and players participated in the surveys (8 0/Cs, 23 players).

Logistics of Data Collection. Data were collected at four sites (Figure 1 and Table 1). ARI and Human Resources Research Organization (HumRRO) scientists collected BN Task Force data at Ft. Knox. NAWCTSD scientists collected FAC-A and OV10 Pilot data at NAWC, Patuxent River, and MULE-Team data at San Diego (Year 1) and Mesa (Year 2). Scientists and one military officer from Armstrong Laboratory collected F16 Pilot and O/C data at Mesa.

Table 1. Logistics of Data Collection for 1994 and 1995 SIMEXs.

SITE	PARTICIPANTS	PROCEDURES
Ft. Knox	BN Staff, ALO Company Team, O/Cs	Two data collectors (DCs) administered surveys, conducted interviews, and kept activity logs
San Diego '94 Mesa '95	MULE Team & O/Cs	One DC administered surveys, conducted interviews, kept activity logs
Mesa	F16 Pilots & O/Cs	Two DCs administered surveys, conducted interviews, kept activity logs
Patuxent River	AFACs, OV10 Pilots, & O/Cs	One DC gave surveys, conducted interviews, kept activity logs

Data Collection Instruments

To develop the assessment instruments, we drew on prior surveys of CAS-DIS training (Holstead, 1989), ARI survey research (Babbitt & Nystrom, 1989a, 1989b), and other sources (e.g., Converse & Presser, 1986; Rea & Parker, 1992). Since then new scholarship has enriched knowledge about survey methodology (Kraut, 1996; Schuman & Presser, 1996).

Background and Experience Questionnaire. This identified the respondent's role in the MDT2 exercises, and tapped demographic information and CAS experience. It addressed institutional and field training experience. Three versions were used, one for ground experience (Army and MULE elements) and two for air (F16 and AFAC elements). The background and experience data helped us track and process survey and responses.

Initial Training Value Questionnaire. This was designed to tap initial expectations about the ensuing week's training, and reactions to specific features of MDT2 (Appendix A). It contained 11 statements and solicited responses on a 6-point, agree-disagree scale. Nystrom and Babbitt (1989a, 1989b) and Converse and Presser (1986) express reservations about the use of an indifference point (neither agree nor disagree). Consequently, we didn't use one.

The Training Value Questionnaire content, was developed, in part, using methodology described in Babbit and Nystrom (1989). As suggested there, we conducted a number of brain storming sessions with the MDT2 multi-Service project staff to generate a list of characteristics reflecting anticipated 'value' characteristics of MDT2. This list was used to generate a set of candidate statements for inclusion in an Training Value Questionnaire. We then selected groups of two or three statements which appeared to cluster under the value dimensions of our assessment model.

Final Training Value Questionnaire. This final survey repeated the initial Training Value Questionnaire items. But, we added an item asking the respondents to estimate transfer of training effects. This 'final' Training Value Questionnaire was administered at the end of training in 1994 and 1995. In the 1994 administration, we asked for marginal comments. But in 1995 we provided lined space for comments after each survey item. This small change had a dramatic effect. The number of comments increased from 27 in 1994 to 175 in 1995. If this format change were proven to be reliable, it would add a modest but useful specification to survey design. Babbitt and Nystrom (1989) address the use of comments in check-lists only briefly and tangentially[6]

Training Objectives Survey. This instrument (Appendix A) assessed how well CAS TOs were represented in MDT2 and how important it was to provide training on those objectives in a multi-Service environment (Question 2A). The MDT2 exercises were based on 25 CAS TOs[7], derived from doctrinal literature (Department of Defense, 1996). The descriptive titles of these TOs are listed in Table 1. In 1994 the Training Objectives Survey included 17 out of the original 25 TO's. The 17 selected were judged most clearly applicable to multi-Service training. In 1995, however, we chose to use all 25 TO's.[8] Respondents were asked to estimate how well those objectives were represented in the MDT2 distributed environment. Their responses were viewed as estimates of content validity of that environment. These, in turn, were viewed as indirect measures of training effectiveness. Separate versions of this questionnaire were administered to:

- Battalion staff and Air Liaison Officer (ALO) at Ft. Knox,
- Company Commander, Executive Officer, Fire Support Team Leader (FIST), and Scout Platoon Leader at Ft. Knox, KY,
- O/Cs at Ft. Knox
- Forward Air Controller - Airborne (FAC-A) and his observer/controller at Patuxent River, MD,
- CAS pilots and their controller(s) at Armstrong Laboratory, Mesa, AZ, and
- Laser designator teams and O/Cs at San Diego, CA (1994) and Mesa, AZ (1995).

The TO Survey is an empirical version of Burnside's analytic method for estimating the content validity of simulation networked (SIMNET) tactical training. Burnside developed check lists for military subject matter experts to assess which company and battalion tasks could be trained on SIMNET. This method was later included in a guidebook for designing virtual simulation training exercises (Hoffman, Graves, Koger, Flynn, & Sever, 1995).[9] It was also included in the Simulation Networking, Training Requirements Relational Data Base (Meliza, 1993)

The TO Survey also bears some resemblance to the 'backward transfer paradigm' used by ARI - Ft Rucker. In this latter paradigm participants are asked to estimate how closely specific simulator characteristics resemble those of actual equipment (Stewart, 1994). A backward paradigm questionnaire per se was not used for MDT2 assessment. (However, the group interviews did ask open-ended questions about similarities and differences with other training contexts.)

Post-Exercise Interviews. On the final day of all the exercise in 1994 and again in 1995, we conducted group interviews. These interview were designed to elicit strengths and weaknesses of the MDT2 training system and provide additional insights from exchanges of views among the participants. In addition they were expected to highlight patterns of opinion across echelons, sites, Services, and types of participants (players vs. O/Cs). See Eyre (1994, P. 102) for an illustration of "pattern extraction" from interviews.

The interview protocols used in 1994 were revised for 1995 to include questions on the after-action review (Appendix A). In MDT2-94, the interviews provided supplementary data on value-added (i.e., Question 1). In MDT2-95, they emphasized Questions 2a (coverage of objectives) and 2b (training methods and management).

Table 2. Training Objectives (TOs) for MDT2 Exercises (Shaded areas designate key objectives for special training emphasis in the 1995 exercises).[10]

Number	Title
1	Determine battalion mission intent and concept of operations
2	Determine the enemy situation
3	Develop CAS target priorities
4	Develop priority of intelligence collection assets to detect CAS targets
5.	Integrate CAS and other fire support elements with maneuver actions
6	Institute fire support control/coordination measures
7	Initiate airspace coordination measures (ACAs)
8	Incorporate SEAD in the fire plan
9	Protect laser team
10	Prepare a decision synchronization matrix
11	Establish methods to identify targets during CAS operations
12	Establish methods to identify friendly troops during CAS operations
13	Conduct a fire support/CAS rehearsal
14	Pass preplanned CAS targets to higher headquarters
15	Prioritize all CAS requests from subordinate commanders
16	Pass immediate targets and on-call target updates to higher headquarters
17	Provide initial brief to pilots and controllers
18	Update airborne pilots as necessary
19	Perform communications check among all fire support and CAS participants
20	Control CAS air attack
21	Confirm status of friendly air defense
22	Arrive on station and establish initial communications
23	Synchronize CAS attack with other direct and indirect fires
24	Conduct CAS attack
25	Return from and assess CAS

QUESTION 1: WHAT TRAINING VALUE DOES MDT2 ADD?
Results

Data for the five dimensions of value include the initial and final Training Value Questionnaires in 1994 and a final Training Value Questionnaire in 1995. Questionnaire comments and group interview testimony supplemented the check-list data. Table 3 summarizes results of the initial and final Training Value Questionnaires for May 1994. This table lists the five value dimensions, related questionnaire items, and percentage of participants who agreed with each item. The table shows high expectations for all five dimensions. It also shows that these high expectations were maintained across the training week. The lowest percentages are for 'credibility'. Even these range from 74% to 94% agreement.

Table 4 summarizes comparative data for the final (i.e., post training) Training Value Questionnaires in 1994 and 1995. Agreements about the value of MDT2 training remain consistently high across the years and across value categories, except for 'credibility'. Note, the substantial drop in percent agreement for credibility items. Percent agreement for 'realism of CAS feedback', for example, dropped from 94% to 72%. Table 5 presents representative comments for each of the training value categories. These comments are drawn from survey and interview data in 1994 and 1995. Note the critical comment for 'Credibility' in 1995. This is typical of comments expressing concerns about simulator limitations on exercise realism. (A comprehensive analysis of interview protocols from the 1995 exercises is available upon request. As part of the analysis, statements related to the value dimensions were extracted across sites and grouped by value dimension.)

Tables 6 and 7 present opinion data by site and value dimension. They show percent of agreement and relative frequencies of agreement 'averaged' across items. These tables are presented to shed light on similarities or differences across the Services in ratings of MDT2 training value. Inspection of the tables shows that, for 1994, the total positive response rate was at least 90% for three of the sites: Knox, Mesa, and Patuxent River. For 1995 the total positive response dropped substantially for Knox and Patuxent River, increased substantially for the DF-MULE site, and remained uniformly high for the Air Force site. Note especially large drops in 'credibility' judgments at Knox and Patuxent River. Table 8 provides a further break-out of the data from Knox. Responses of Company Team have been separated. These, along with interview data, indicate that Company Team responses account for much of the decline in value ratings at Knox from 1994 to 1995.

Table 3. Initial And Final Training Value Questionnaire - 1994 Sample

ISSUE	SURVEY ITEM (31 Subjects across all sites)	% AGREE* Initial	Final
1a. Need	12. The opportunity provided by MDT2 to practice with personnel from other Services is necessary for training CAS	97	90
	13. MDT2 is a good training system for CAS because it focuses on critical training needs	90	90
	18. Given the chance, I would like to train with MDT2 on a periodic basis.	90	94
1b. Credibility	14. MDT2 can be an effective trainer for CAS with only a few minor modifications	74	81
	15. A positive aspect of MDT2 is that it gives more realistic feedback on CAS kills than in field exercises or at Combat Training Centers	88	94
	16. I can apply more realistic CAS tactics in MDT2 than I can in field exercises or at CTCs.	77	77
1c. Multi-Service Value	7. Experience on MDT2 will make me better able to interact with members of other Services to plan for and execute CAS missions in combat	93	90
	8. Training with MDT2 will give me a better understanding of the jobs and roles of personnel from other Services in planning and conducting CAS	84	84
1d. Role in Training Cycle	6. Experience on MDT2 will better prepare me for field exercises on CAS missions, such as those at Air Warrior and NTC	93	87
	11. Training on MDT2 can supplement Service-specific CAS training	91	87
1e. Expected Impact	10. The training that MDT2 provides can be applied directly to combat	97	97
	22. Estimate the extent to which your experience with MDT2 has affected your ability to perform your role in a mission that uses CAS	NA**	93#

* % Agree = Percentage of troops who checked slightly, moderately, or strongly agree.
** NA: Not applicable. This statement was used only in the final survey.
# Choices were: no change, slight increase, moderate increase, large increase, extreme increase in combat effectiveness.
93% of the troops marked slight to extreme increase. 7% marked no change.

Table 4. Final Training Value Questionnaire (1994 Vs. 1995)

ISSUE	SURVEY ITEM (Comparisons of Year 1 and 2; 31 subjects each year, across all sites.)	% AGREE* 1994	1995
1a. Need	12. The opportunity provided by MDT2 to practice with personnel from other Services is necessary for training CAS	90	90
	13. MDT2 is a good training system for CAS because it focuses on critical training needs	90	74
	18. Given the chance, I would like to train with MDT2 on a periodic basis.	94	83
1b. Credibility	14. MDT2 can be an effective trainer for CAS with only a few minor modifications	81	55
	15. A positive aspect of MDT2 is that it gives more realistic feedback on CAS kills than in field exercises or at Combat Training Centers	94	72
	16. I can apply more realistic CAS tactics in MDT2 than I can in field exercises or at CTCs.	77	53
1c. Multi-Service Value	7. Experience on MDT2 will make me better able to interact with members of other Services to plan for and execute CAS missions in combat	90	90
	8. Training with MDT2 will give me a better understanding of the jobs and roles of personnel from other Services in planning and conducting CAS	84	89
1d. Role in Training Cycle	6. Experience on MDT2 will better prepare me for field exercises on CAS missions, such as those at Air Warrior and NTC	87	90
	11. Training on MDT2 can supplement Service-specific CAS training	87	77
1e. Expected	10. The training that MDT2 provides can be applied directly to combat	97	100
impact	22. Estimate the extent to which your experience with MDT2 has affected your ability to perform your role in a mission that uses CAS	93#	94#

* % Agree = Percentage of troops who checked slightly, moderately, or strongly agree.
# Choices were: no change, slight increase, moderate increase, large increase, extreme increase in combat effectiveness.
In 1994 93% of the troops marked slight to extreme increase. In 1995 the figure was 94%

Table 5. Sample Of Comments By MDT2 Participants

ISSUE	Comments
Need	"Great training concept. We only work with the air crews 1-2 times annually, yet it is required of me, in my job, to be proficient in calling in CAS." ('94) "We have other simulations that allow us to do the staff training. This is the only one that allows us to do CAS." ('95)
Credibility	"Real time, free play, fast paced environment provided realism. Generates 'the fog of war' in safe environment" ('94) "The game was to see 3.5 clicks out ... In a desert environment that doesn't allow very much survivability for us. ... If we can get it farther out it would be more realistic for us" ('95)
Multi-Service Value	Comment about the objective 'Initiate Airspace Coordination Areas' "This is a highly complex task, with 'perishable' experience. MDT2 provided excellent training in this area. Looks good. Keep it coming!" ('94) "There was great interaction between the Air Force, Marines, and the Army. I think this is a great setting to allow that interaction to happen." ('95)
Role in Training Cycle	"Every two months for my personnel. It is very necessary to have all Services involved." ('94) "I would try to use it repetitively, say on a quarterly basis as a refresher training for command & control and keeping the staff proficient and for also training new staff members. We go through people ... personnel turnover is incredible." ('95)
Expected Impact	"TREMENDOUS POTENTIAL! Should help to greatly reduce instances of fratricide. Let's get it fixed!!" ('94) "I feel very confident now at the end of the week, that I could go to the NTC world and integrate CAS effectively." ('95)

Table 6. Summary of Opinions by Value Category Across Sites Final Training Value Questionnaire - 1994

VALUE	KNOX	MESA-AF	DF-MULE*	PATUXENT
Need	95% (63/66)	92% (11/12)	67% (8/12)	100% (3/3)
Credibility	86% (57/66)	92% (11/12)	67% (8/12)	67% (2/3)
Multi-Service	93% (41/44)	100% (8/8)	38% (3/8)	100% (3/3)
Role in Cycle	86% (38/44)	100% (7/7)	75% (6/8)	100% (2/2)
Impact	100% (22/22)	100% (4/4)	75% (3/4)	100% (1/1)
Total	91% (221/242)	95% (41/43)	58% (28/44)	92% (11/12)

*located at the Naval Personnel Research and Development Command, San Diego, CA

Table 7. Summary of Opinions by Value Category Across Sites, Final Training Value Questionnaire - 1995

VALUE	KNOX	MESA-AF	DF-MULE#	PATUXENT
Need	75% (39/52)	100% (15/15)	92% (11/12)	83% (10/12)
Credibility	57% (29/51)	87% (13/15)	67% (8/12)	33% (4/12)
Multi-Service	84% (27/32)	100% (10/10)	88% (7/8)	100% (8/8)
Role in Cycle	79% (26/33)	100% (10/10)	100% (8/8)	63% (5/8)
Impact	100% (17/17)	100% (5/5)	100% (4/4)	100% (4/4)
Total	75% (138/185)	96% (53/55)	86% (38/44)	70% (31/44)

#Collocated with Air Force at Armstrong Laboratory, Mesa, AZ

Table 8. Summary of Opinions by Value Category at Knox, Final Training Value Questionnaire - 1995

VALUE	Knox	Knox Minus CoTm	Company Team
Need	75% (39/52)	78% (31/40)	66% (8/12)
Credibility	57% (29/51)	69% (27/39)	16% (2/12)
Multi-Service	84% (27/32)	96% (23/24)	50% (4/8)
Role in Cycle	79% (26/33)	92% (23/25)	38% (3/8)
Impact	100% (17/17)	100% (13/13)	100% (4/4)
Total	75% (138/185)	83% (117/141)	45% (21/44)

QUESTION 1: TRAINING VALUE
Discussion

General Trends in Perception of Value

The results summarized (i.e., 'rolled-up) across Services show that the MDT2 participants saw value-added along each of the value dimensions for 1994 and 1995. But the exceptions to this generalization are compelling and instructive for further developments and uses of the MDT2 technology. For example, in 1994, the participants from Knox, Mesa-AF, and Pautuxent River agreed after initial exposure that MS-DIS was essential for training CAS. And, the consensus was maintained to the end of training. However, he MULE-DF participants in San Diego gave MDT2 substantially lower 'grades' for all five value dimensions. A plausible reason is that the MULE site was frequently disconnected from the net because of technical failures and therefore was frequently excluded from portions of exercises.

This problem was solved in 1995 when DF-MULE collocated with the Air Force at Williams. Moreover, the participants indicated that they liked being collocated with the Air Force since collocation facilitated communication and allowed the MULE participants to learn, first hand, about Air Force operations. The value ratings rose dramatically. Even here though there was an exception. Credibility ratings were low for both sites (San Diego in 1994; Mesa in 1995). The MULE simulation had a number of unavoidable (for technical reasons) artificialities which existed at both sites.

The generally high 1994 ratings at Knox dropped substantially in 1995 . The lower valuation may reflect a difference between reserve and active army perspectives and needs. Whether such differences are reliable would require further study. In addition, the 1995 survey and interview data, surfaced a particular problem at the Company Team echelon. These players felt that their role in CAS was marginal and therefore they did not benefit from the training - that they were functioning as 'training aids'. They suggested replacing the live Company participants with semi-automated forces. Nonetheless they placed high value on MDT2 for those involved in CAS. Note, for example, (Table B-3) that their rating was 100% for Item 12 ("... MDT2 ... is necessary for training CAS"), but only 25% for Item 18 ("... I would like to train with MDT2 ..."). Their interview statements reinforce these numbers.

The disparate numbers coupled with interview comments point up a methodological issue, given the complex, multi-echelon, multi-Service DIS environment. Items may need to accommodate for conditional responses, e.g., "MDT2 satisfies a training need, but not at my echelon."

Where you can anticipate such conditional reactions, write the survey item accordingly. Otherwise, encourage respondents to explain conditional responses in comments. This problem is not unique to DIS, but it would be aggravated by the DIS environment, where varied populations of trainees are likely.

Specific Value Issues

Need for Multi-Service Distributed Interactive Simulation (MS-DIS).

The survey numbers indicated generally that participants saw value-added in the training because of training need satisfaction. Their interview data suggest that need is satisfied (or limited) in a number of ways:

1. There is no alternative to practicing planning/decision making for CAS and then realistically seeing the impact of those decisions.

2. There is no feasible way to practice CAS frequently and repeatedly. Field operations are more realistic in some respects but are expensive and may be dangerous. For example, trainees at the Air-Ground Operations School (AGOS), Hurlburt Field, FL can control the final approach of aircraft accurately because they can see the plane's attitude and nose direction. MDT2 lacks the necessary visual fidelity for accurate final control.

On the other hand, trainees can only practice a few times in the field because each practice run is expensive and time consuming. Moreover, aircraft are not always available and dry runs may have to be substituted. The value of the field setting is diminished under such circumstances. Similar constraints may apply to any field training setting.

3. Currently there are too few opportunities for the Services to practice CAS with each other.

4. MDT2 provides a way to concentrate on a phase of battle that otherwise might be lost in larger-scale exercises. Participants mentioned use of MDT2 as 'ramp-up' for NTC rotations. Coincidentally, analyses by the Air-Ground Operations School indicate that CAS is not well trained at the National and Joint Training Readiness Centers (Center for Army Lessons Learned, 1995). At CTCs CAS is a very 'little fish in a big pond'.

5. Multi and Joint Training are the 'wave of the future.'

6. Notwithstanding all the above, survey results and analyses - especially from company team players - indicated that need satisfaction is ***relative*** and ***fragile***. If this principle is not considered in assessing training need, misleading data and conclusions are possible.

Credibility and Realism. Year 1 participants indicated that MDT2 training has very high potential for credibility and realism. Initial expectations about credibility and realism were sustained to the end of the training week. They identified many problems which need to be fixed before that potential is reached. But their interview comments suggest that they perceive the problems to be technical, manageable, and surmountable.

The credibility ratings for Year 2 decreased substantially. Examination of data across sites indicated that much of the decrease in credibility was attributable to Ft. Knox (Active Army) and Patuxent River (Airborne Forward Air Controller) participants. Table B-3 shows a disproportionate contribution by the Company Team Participants to the decrease at Knox.

Comments by the Knox participants revealed shortfalls in tactical realism, e.g., unrealistic operations orders, intelligence information and planning, maneuver areas, detection ranges, casualties, threat capability, and use of scouts, as well as inaccurate representation of some TOs underlying the exercises.

Problems at Patuxent River seemed related to changes from a cockpit simulator in 1994 to a work station simulator and helmet mounted display in 1995. Participants identified distortions of tactical realism related to characteristics of the simulators.

These problems did not emerge until after the week of exercises in 1995. A lesson to be learned: for any distributed simulation, management control measures are needed to detect site problems in advance or at least through assessment procedures after the fact.

Yet, the interview comments, especially by O/Cs across the sites, indicate: 1) an appreciation that training simulation is not actual combat and 2) a tolerance for less than full fidelity. Their comments generally reflected fidelity characteristics that could be improved, e.g., more realistic communication, greater visual ranges, and better translation of TOs into scenario design.

Air Force participants provided the highest credibility ratings. These were reflected in interview comments which favorably compared MDT2 with Air Warrior, a joint exercise with the Army at NTC. Participant's noted similarities of scenario and problems.

Players differed in several ways from O/C's with interesting consistency across the Services on the Credibility dimension. The interview comments of the O/Cs indicated greater tolerance of simulation limitations (SIMLIMs). In some cases they opined that the SIMLIMs didn't matter because the limitations did not interfere with the key TOs. In other cases they indicated that the SIMLIMs were just engineering problems that were surmountable or could be 'worked around'.

Among the players, the Company Team (CoTm) participants were the most critical of credibility. This showed in the survey data and was reinforced by the interview comments. But even here, the criticisms were constructive. Their comments more than those of any other participants pointed to mistranslations of objectives into scenario design, rather than defects in the underlying MDT2 training concept. In effect their comments said: "you need to do a better job of matching SIMEX design to TOs". For example, eliminate live CoTm players, since they are not involved in close air support and serve only as training devices. Replace them with SAFOR elements.

Reactions suggesting ineffective translation of TOs point to a need for a more systematic and concerted effort to check the scenario design against TOs. This could be done through 'bootstrapping': write the TO, write the scenario, and then check its characteristics against the TO.

Multi-Service Value. The 1994 participants (except for DF-MULE players)started with very strong expectations that MS-DIS would improve their skills at interacting with other Services and their knowledge of the roles played by the other Services. These expectations were sustained from the first to the last exercise.

Ratings by DF-MULE participants were notably low. Their negative reactions are plausibly related to severe networking problems. San Diego was frequently off the net because of technical difficulties.

Positive ratings continued into 1995, including a dramatic increase in ratings from the MULE-DF participants. This improvement is plausibly related to a site change. In 1995, the DF-MULE was moved to Armstrong Laboratory in Mesa. Network problems were negligible and Marine Corps players were able to interact face to face with F16 pilots (between, not during exercises).

There *was* a modest decrease in the Fort Knox ratings. But these were traceable to CoTm responses on Question 7. Question 7 ("...MDT2 will make me better able to ... plan for and execute CAS...") addresses a role that the Team did not play. Interview comments are consistent with the low CoTm responses to Question 7.

The interview data indicated a number of ways in which participants see multi-Service value::

(1) Trainees learn other-Service nomenclature and procedures. This learning improves ability to communicate and to coordinate.

(2) They make and learn to avoid mistakes that would otherwise be made if the Services came together for the first time in battle.

(3) They accomplish the above at a relatively modest cost, e.g., by increasing the efficiency of follow-on operational tempo (OPTEMPO) training.

(4) The teleconferenced AAR reinforces the multi-Service training value even where feedback is not directly aimed at particular players: "Value is added on learning, having displayed and discussed everybody's role".

Role in Training Cycle. The MDT2 participants showed a strong consensus that MS-DIS could supplement their Service specific CAS training and prepare them for training at combat Service centers. Again the initial expectations were sustained to the end of training. 1995 ratings of the value of MDT2 for CTC 'ramp-up' were comparably strong. But ratings decreased notably for the value of MDT2 as a supplement to Service specific training. This decrease is accounted for primarily by Company Team responses. It is consistent with interview comments indicating that MDT2, with its emphasis on CAS, did not provide useful training at the Company Team level. Hence, there was nothing to supplement.

The responses of the Patuxent River (AFAC) participants also dropped - substantially. Most of this decline is traceable to Question 6. Most likely, it is an artifact of an imprecise question. Question 6 references NTC and Air Warrior, neither of which - we believe - is attended by Patuxent personnel. More likely they do OPTEMPO at the Marine Corps field training facility, '29 Palms', near Barstow, California. ***Lesson learned for multi-site, multi-group training: examine assessment questions carefully to insure that they are tailored to each of the sub populations involved in the training.***

The interview data offer many useful ideas about the role of this kind of training in the total 'pipeline'.

(1) Refresher training for command and control staff

(2) Way to counter frequent turnover

(3) Preparation for NTC - 30 to 60 days out

(4) Pre-combat, e.g., rapid mobilization for Bosnia, Somalia, Desert Storm

(5) Component of Blue Flag, especially to validate the computer model or cross check results of SAFOR

(6) Component of Air Ground Operations School training (with higher fidelity displays to allow accurate visual judgments by Air Liaison Officers (ALOs) or Enlisted Tactical Air Controllers (ETACs).

(7) "Come as you are" training. Units with their own simulators at home station can dial up other units for impromptu, 'spur of the moment' part-task drills, e.g., F16s doing drills with AFACs.

Some additional comments were made about how to use MDT2:

(1) Incorporate it into the 'events' list every year.

(2) Watch out for negative transfer, e.g., getting too comfortable where real combat would pose a danger.

(3) Get personnel from the combat training centers involved in designing or reviewing exercises to maximize connection to CTC training.

(4) Make sure that the trainees come with pre-requisite Service-specific skills well developed. i.e., don't violate 'crawl, walk, run' philosophy.

Expected Impact of the training. Players and O/Cs felt that MDT2 training could be applied directly to combat. Moreover, the players felt their skills for CAS supported combat were actually increased as a result of the MDT2 experiences. In 1994, expectations for positive impact of MDT2 were sustained from the initial to final day of training, in spite of the intervening technical problems. A strong consensus extended to the 1995 exercises, because of the MDT2 training.

The interview comments primarily addressed the impact of this type of training on subsequent training events, e.g., NTC or Air Warrior. The most impressive and useful comment was made by one of three F16 pilots who had been to the 1994 MDT2 exercises, had then gone to Air Warrior, and finally had returned to MDT2 in 1995. He felt that these three pilots were several exercises ahead of the other Air Warrior pilots in skill level. The Battalion Staff expressed increased confidence in their ability to deal with CAS at NTC. The Company Team players agreed that this was effective training at the staff level. However, they noted that MDT2 was not effective for them. They felt they served as training devices rather than as trainees.

Value Ratings Across the Services

Three major trends emerge from the site-specific data. The Air Force gave consistently strong training value ratings for the 1994 and the 1995 exercises. Ratings by the Army at Knox and AFACs at Patuxent River, dropped notably from Year 1 to Year 2. Ratings by the Marine Corps MULE participants improved substantially from Year 1 to Year 2.

Reasons for the drop in Army ratings have been mentioned briefly already. Interview data indicated a sensitivity to limitations in tactical realism and imprecise translation of

objectives into battle scenarios, especially at the company-team level. These limitations certainly existed in Year 1, but we think the active unit of Year 2 was less forgiving, than the reserve unit of Year 1. Improvement in rating by the DFO-MULE participants appeared to be related to a change in site from San Diego to Armstrong Laboratory. Along with this came significant technical improvements including better network connections.

Reservations expressed by Patuxent River were related to limitations in their simulation facilities, especially, the work station in Year 2. This replaced a cockpit simulator from Year 1. The problems, which surfaced only after the exercises ended, underscore the need in DIS for system managers and exercise directors to attend carefully to the technical adequacy of each site and do so early in planning. Technical problems at Patuxent River, clearly 'fell through the cracks' and negatively impacted the subsequent MDT2 exercises. This is a lesson to be learned - not a criticism, since the MDT2 program was an R & D effort aimed at discovering new ways of doing multi-Service training business.

Hawley-Christ(HC)Model.

The results are consistent with the HC definition of value added (pages 7 and 8 of this report). MDT2 provides training, not otherwise available. In addition it can provide that training at a tenth the cost of a potential, alternative field exercise (Orlansky et al. 1996). The data also support the HC conclusion that MDT2 satisfies suitability [if not used for basic CAS skill training], and potential effectiveness.

Methodological Challenges for Distributed Training: Training Value Assessment.

Overview. We will address methodological challenges in this and subsequent discussion sections. These challenges have particular impact in DIS environments because of the technology itself or the kinds of likely training: scenario-based, cooperative or joint training across multiple organizations with documentably different populations of trainees.

Challenge 1 Relativity and fragility of the 'value' concept. A challenge for DIS is to summarize a training system's value with a simple index like cost-effectiveness ratio or the measures derived from Training Value Questionnaire frequencies. But achieving this kind of simplicity is problematic when multiple sub-populations within and across organizations come together as they did in MDT2 and as they do in multi- and joint-Service training.

The Training Value Questionnaire data indicated variation in perception of value across sites, echelons, and training events (i.e., years). This variation could obscure summary values unless the data are also examined by sub-population as they were for this research. Such a dissection, however, may have to contend with small frequencies. In a production system, the ideal solution is

to cumulate data per sub-population across many training events, so that respectable frequencies can be reached.

Challenge 2 Clarity and accuracy of the questions. A related challenge is to minimize distortions in the wording of questions preferably by having sites or sub populations pre-view the questions. For example, we found two potential distortions - after the fact, though. The Regular Army (1995) CoTm players saw great value in MDT2 for the Battalion Staff, but very little for their echelon. They could have reflected either perception in their survey responses. But there was no provision, except in the request for comments and interviews, to clarify the distinction between value for one but not another group of players.

A second source of distortion appeared in Question 6 which asked for a comparison with NTC or Warrior Flag. The players at the Patuxent River site had not been to these exercises and hence found some ambiguity in the question. This underscores the need stated earlier to pre-test all survey items with the relevant sub-populations to uncover problems with item structure.

Challenge 3 Interpreting the numbers. Even with well written items and analyses by sub-populations, the meaning of the analyzed frequencies will not be crystal clear unless the participants provide some comments. These are difficult to get because surveys usually come at the end of a long, tiring training event (rotation). Stressing the need for their comments and providing lined space after each question will help.

QUESTION 1: TRAINING VALUE
Conclusions

1. Strong support for the value of MDT2 training was found across the Services, with most consistent support from Air Force participants.

2. The type of training provided by MDT2 is critical and should be implemented. This view was generally shared across the Services.

3. No feasible alternative currently exists to practice planning and decision-making for CAS and then realistically observe the impact of those decisions.

4. MDT2 training was viewed as especially valuable for 'ramp-up' or other part-training support to large scale training exercises.

5. However, players and O/Cs also felt that MDT2 training could be directly to prepare/rehearse for combat.

6. The training system problems which limit MDT2 value can be and should be fixed. Some key solutions include:

- Systematic methods to insure accurate translations of joint TOs into exercise scenarios and feedback measures.

- Management controls to detect site-specific problems which might limit training value at other sites.

- Selected improvements in fidelity. For example, Increases in range of visibility and delectability to more nearly match those available in field conditions.

7. On-going value assessment should be part of any larger training effectiveness assessment.

QUESTION 2A: HOW WELL WERE THE TRAINING OBJECTIVES COVERED BY MDT2?
Results

We have addressed the perceptions of the participants about overall value of MDT2. But what did they think about the training itself? First, were the proper training tasks (objectives) selected, how well were they implemented in the MDT2 system, and how easy or difficult was it to perform those tasks in the MDT2 simulation facilities. Recall that the tasks (i.e., TOs) were selected to meet doctrinal requirements of CAS engagements. Furthermore, for the 1995 exercises a panel of experts in multi-Services CAS operations identified seven key objectives for training emphasis. Still it's important to know whether the troops perceive that the 'right' objectives are driving their training and that those objectives are effectively implemented in the training environment.

These issues were addressed by analyzing data from the Training Objectives Survey. Chi square test statistics were applied to frequency distributions across TOs for questionnaire response categories. In addition, 'importance' and 'training effectiveness' indices were computed for individual TOs. The first set of analyses deals with overall, system-wide perceptions of task importance and training effectiveness. The second deals with perceptions about individual TOs. These analyses bear on the assessment of MDT2. But they also suggest generalizable assessment methodology, particularly for distributed training. We'll address this methodology in the Discussion Section to follow.

Analysis of Summary Frequency Distributions.

Figures 2 and 3 show the summary response distributions to the 17 TOs common to 1994 and 1995 administrations. Table 9 shows overall percentages of 'favorable' responses. Results of a chi square test of the distributions in Figures 2 and 3 are also shown. For 1994, on the importance of additional training, *92.4%* of the responses were favorable (i.e. "Additional training for this task is desirable, highly desirable, or essential). For training effectiveness of MDT2, *73.2%* of the responses were favorable (i.e., provides minimal essential training, more than minimal essential training, or all required training for this task). For 1995, the figures for 'importance' and 'effectiveness' were *84.3%* and *81.7%*. Only the year to year difference for 'importance' was significant. The indications of 'no training' or negative training were less than 6% for both years.

Table 10 summarizes a comparison between frequency data for the key 7 TOs and the remaining 18 of 25 TOs assessed in 1995. The percentages of favorable responses are nearly identical. Chi square for the underlying distributions was nonsignificant, i.e., the distributions were indistinguishable.

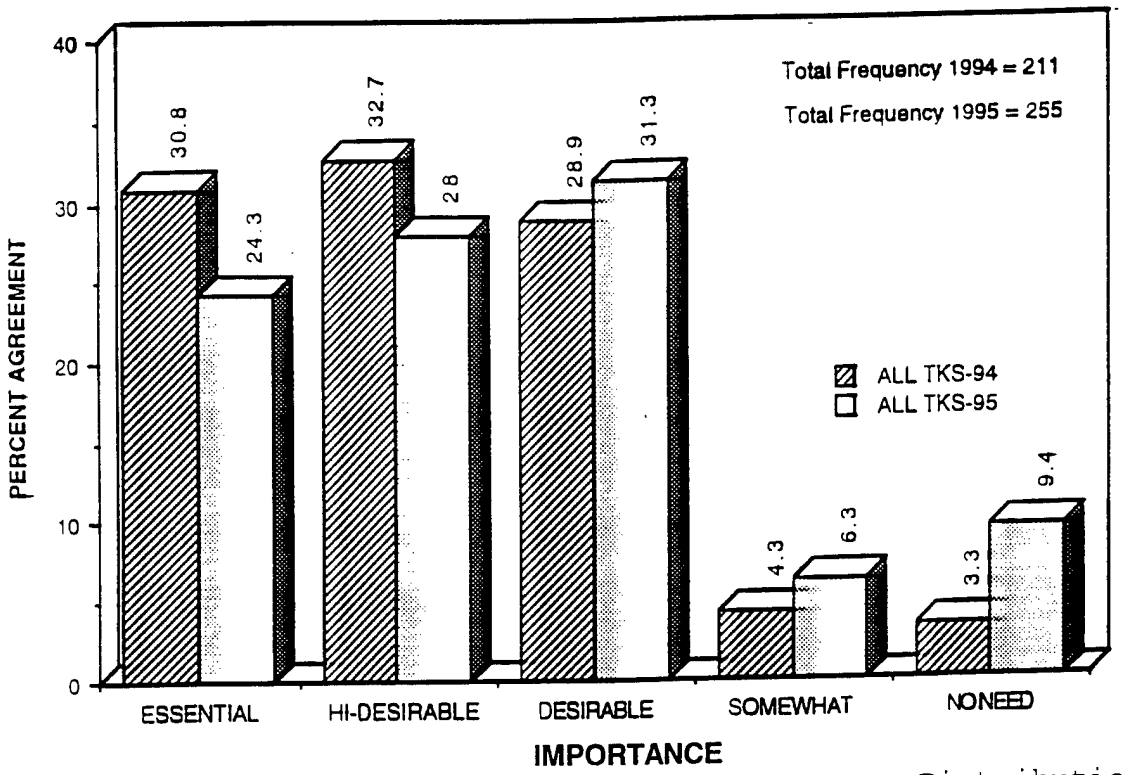

Figure 2. Training Objectives Survey: Summary Distributions for 'Importance.'

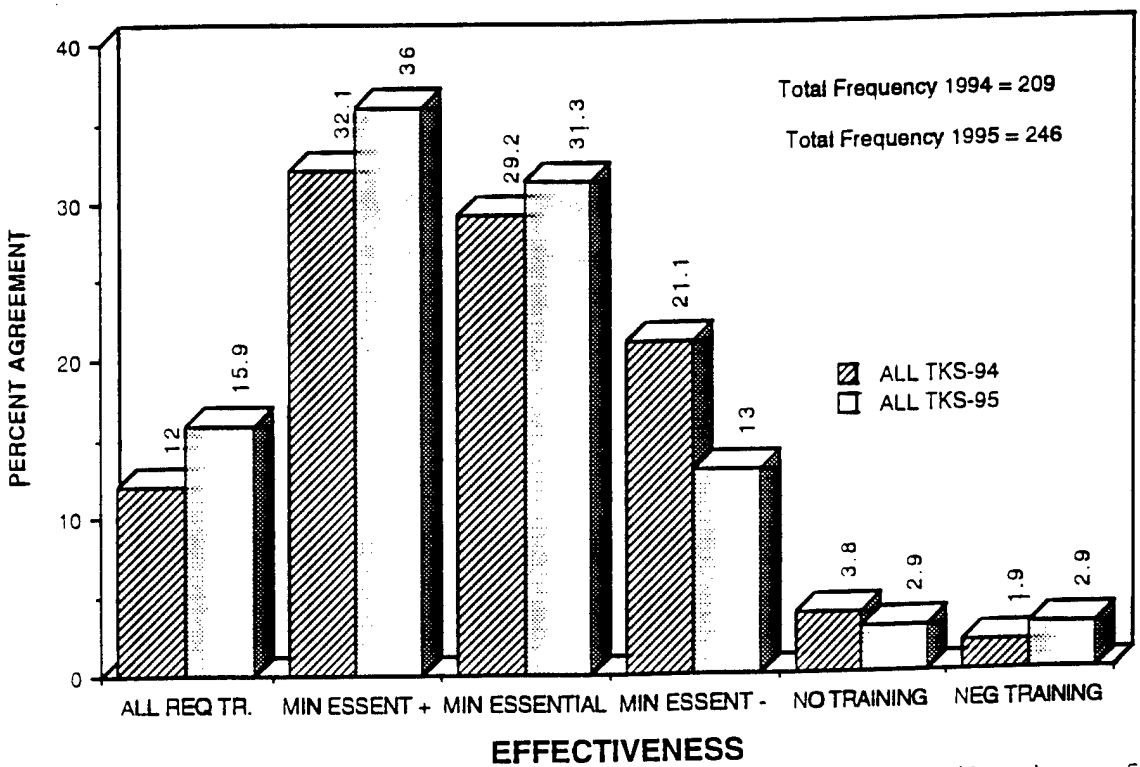

Figure 3. Training Objectives Survey: Summary Distributions for 'Effectiveness.'

Table 9. Percentages Of "Favorable" Training Objectives Survey Responses For 17 TOs Common To 1994 and 1995

RESPONSE	MAY 94	FEB 95	$\chi 2$, P ≤
Additional training is desirable, highly desirable, or essential	92.4	84.3	.0411#
Provides minimal essential, more than minimal, or all training required	73.2	81.7	N.S.#
Provides no training or negative training	5.7	5.8	

Chi square tests were performed on the frequency distributions shown in Figures 2 and 3.

Table 10. Percentages Of Favorable Training Objectives Survey Responses For Non-Key And Key TOs In 1995.

RESPONSE	18 Non Key TOs	7 Key TOs	$\chi 2$, P ≤
Additional training is desirable, highly desirable, or essential	82.8	82.1	N.S.
Provides minimal essential, more than minimal, or all training required	81.7	81.6	N.S.
Provides no training or negative training	5.8	3.9	

Computation of Derived Importance and Effectiveness Ratings for Individual TOs.

Numerical values were assigned to survey categories for the 'importance' and 'training effectiveness' scales for each of 17 tasks surveyed in 1994 and 25 tasks in 1995.[11] Then the mean 'importance' and 'effectiveness' scores were multiplied for each TO to yield a 'training value' index. The results of these computations for 1994 are presented in Tables 11 ranked ordered for 'importance'. Key TOs are highlighted to show how they fit into the ranking. Similar data are presented for 1995 in Table 12.

Note that the Key TOs in Tables 11 and 12 are distributed throughout the list rather than concentrated at the top where one would expect them to be if participants agreed with MDT2 staff on what the most important TOs are for additional training in MDT2. This disparity will be addressed in the Discussion.

Another view of year to year contrasts was generated from correlations for 'importance' and 'effectiveness'. A correlation (1994 vs. 1995) of .412 for 'importance' was significant ($n = 17$, $p < 0.05$). The correlation for 'effectiveness' was not significant.

Table 11. Rank Order of Training Objectives by Importance - 1994

TRAINING OBJECTIVE	NUMBER OF RESPONSES	***IMPORTANCE (i)***	EFFECTIVE. (E)	TRAINING VALUE (IxE)
24. Conduct CAS attack	17	***3.4***	2.3	7.82
13. Conduct CAS rehearsal	8	***3.2***	2.2	7.04
16. Pass immediate targets	5	***3.2***	1.2	3.84
12. Methods to ID friendly	14	***3.1***	1.6	4.96
7. Initiate airspace areas	9	***3.0***	2.4	7.20
9. Protect laser team	8	***3.0***	1.4	4.20
8. Incorporate SEAD	18	***2.9***	2.7	7.83
25. Return from/assess CAS	13	***2.9***	2.0	5.80
17. Provide initial brief	10	***2.9***	2.3	6.67
22. Arrive on station	19	***2.8***	2.7	7.56
18. Update airborne pilots	13	***2.8***	2.5	7.00
15. Prioritize CAS requests	7	***2.7***	2.1	5.67
21. Confirm status friendly	9	***2.7***	2.1	5.67
11. Methods to ID targets	17	***2.6***	2.1	5.48
20. Control CAS air attack	14	***2.6***	2.3	5.98
1. Determine mission intent	13	***2.5***	2.2	5.50
19. Perform commo check	19	***2.1***	2.2	4.62

Table 12. Rank Order of Training Objectives by Importance - 1995

TRAINING OJBECTIVE	NUMBER OF RESPONSES	***IMPORTANCE (I)***	EFFECT. (E)	VALUE (ExI)
16. Pass immediate targets	4	***3.25***	3.33	10.82
12. Methods to ID friendly	13	***3.08***	1.62	4.99
17. Provide initial brief	12	***3.00***	1.83	5.49
21. Confirm friendly status	8	***3.00***	2.14	6.42
24. Conduct CAS attack	16	***3.00***	2.60	7.8
25. Return from & assess CAS	4	***3.00***	2.92	8.76
20. Control CAS attack	21	***2.62***	2.45	6.42
7. Initiate airspace areas	15	***2.60***	2.00	5.2
8. Incorporate SEAD	17	***2.53***	2.29	5.79
22. Arrive on station	24	***2.50***	2.92	7.3
18. Update airborne pilots	19	***2.47***	2.71	6.69
6. Institute fire spt measures	11	***2.45***	2.18	5.34
23. Synch CAS with other fires	11	***2.45***	2.55	6.25
15. Prioritize CAS requests	3	***2.33***	2.67	6.22
11. Methods to ID targets	19	***2.32***	2.89	6.7
19. Perform commo check	28	***2.32***	2.64	6.12
2. Determine enemy situation	9	***2.22***	1.56	3.46
9. Protect laser team	13	***2.15***	1.25	2.69
14. Pass preplanned - higher HQ	7	***2.14***	2.00	4.28
3. Develop CAS tgt priorities	13	***2.08***	2.00	4.16
5. Integrate CAS with maneuver	12	***2.08***	2.92	6.07
10. Prep. decision synch matrix	11	***2.00***	2.73	5.46
13. Conduct CAS rehearsal	18	***1.89***	2.67	5.05
4. Dev prty intel assets	8	***1.88***	2.00	3.76
1. Determine mission intent	13	***1.85***	1.83	3.39

Use of Training Objectives Survey Data in Instructional Design.

An examination of the frequency distributions for individual tasks may reveal information useful for instructional design or management. It provides a way to 'red-flag' TOs that have not been adequately translated into training scenario design or strategy and to assess whether the training system has improved from application to application. For example, results of an examination of the 1994 and 1995 frequency distributions for 'training effectiveness' are summarized in Table 13. The table shows tasks 'red-flagged' as sources of training problems (rating < 2.00; less than minimal required training). The data show marked improvements in implementation (i.e., translation) for three (shaded) of five TOs identified as problems in 1994.

Table 13. Problem Tasks in 1994 ('effectiveness' rating < 2.00) Compared with 1995 Tasks

Task	Rating 1994	Rating 1995
Protect laser team	1.38	1.25
Methods to identify friendly forces	1.69	1.62
Conduct CAS rehearsal	1.86	2.67
Pass immediate targets	1.50	3.33
Return from and assess CAS	1.53	2.92

The improvements did not actually result from the analysis in Table 13 because that analysis was not available in time for the preparations for the 1995 exercises. But the data are presented here to illustrate a methodology that might be used as part of a fielding of the MDT2 system or any other distributed training system. The methodology would begin with data similar to those in Table 13 to red-flag training system problems.

Solutions would begin with an examination of survey comments and interview data. For example, a player who responded negatively to "Protect laser team" observed the following: "Laser team could not move with protecting element, but instead jumped from OP to OP, usually in front of and ahead of friendly forces. Did not train for survivability." Another player commented: "No direct commo link to MULE team by TF Cdr or staff or FSO". A negative comment on "Return from and assess CAS mission" was annotated with "Pilot had trouble giving back BDA." Here then are indications of inadequate translations of TOs into scenario and potential prescriptions for improvement.

It happens that the deficiencies in these particular cases resulted primarily from simulation limitations (SIMLIM), not oversights in training design, but the comments are still useful.

If neither technology 'fixes' nor 'work arounds' are feasible, then the TOs should be excised from the next training rotation, with an explicit recognition that, while important for a CAS mission, they are not trainable in this particular system.

The next example suggests a problem other than a SIMLIM. A negative response to "Establish methods to identify friendly troops during CAS operations" was followed by "This issue was not briefed or discussed." The comment suggests a failure to explicitly 'design' the TO into the scenario.

The second step would be to re-examine the detailed description of the TOs and to review the exercise protocols to determine how the TOs were implemented in the exercises. The comments above would be now be brought to bear. This review could lead to any number of conclusions: e.g., improve the simulation technology, revise exercise design, train the TO by alternative means, or simply ignore the problem because improvements are not feasible or cost-effective, or the TO has a low training priority.[12] In any case TOs which can't be implemented should be put on a separate list of objectives which are essential to combat but need to be trained in some other type of training event.

Finally, follow-up measures could identify training system improvements across repeated applications and revisions of MDT2.[13] A summary measure of system effectiveness might be indexed by the number of TOs rated 2.00 or more (minimally trained or better). Such comparisons would be problematic for the MDT2 program because of the many confounding sources of variance across years. However, in a fielded, relatively stable version of the MDT2 Training System, comparisons (including summary indices) could be useful as part of an on-going quality control program.

QUESTION 2A: HOW WELL WERE TRAINING OBJECTIVES COVERED BY MDT2? Discussion

Overall Perceptions About Training Value and Effectiveness.

Results show that participants generally believed that the tasks ***(TOs) were satisfactorily implemented by the MDT2 system.*** They also believed that those tasks ***were important to include in training beyond that provided by the Services.*** Even though The MDT2 training system proved to be more a "brassboard" than a prototype. the degree of favorable reaction was respectable. The amount of unfavorable reaction was very small, with substantial opportunity for growth in those tasks judged to be marginally trained. Year to year comparisons, in fact, suggest an increase in perceived effectiveness. On the other hand, perceived task 'importance' decreased, though it was still at a high level. These comparisons are plausible given the technical improvements in 1995 and the change from a reserve Army unit in 1994 to an active one in 1995.

Participant Perceptions of Relative Importance of Key Tasks.

Frequency analysis of the key TOs for 1995 showed that the ***participants did not perceive these tasks to be more important than the other tasks for additional training.*** This finding suggests a disagreement with the MDT2 subject matter experts (SMES) who selected the key TOs. The finding is reinforced by the rank-ordered data on derived task 'importance' ratings. In neither 1994 or 1995 are the key tasks clustered at the top of the ranked lists where one would have expected them, if ***participants*** viewed these as key tasks.

However, the disagreement may be more apparent than real. The SMEs were charged with identifying a small set of the most combat critical inter-Service tasks. The participants were asked to rate the importance of ***additional*** training - not the same question. But the apparent discrepancy does point to the need for a very clear understanding about how key training tasks have been or should be identified.

Generalizable Uses of Training Objectives Survey Data for Training Design and Management.

Key Task Selection. In a production version of the MDT2 system or any distributed training system, the unit commanders and their staffs would set priorities for selecting training objectives and key TOs. Their major criteria would be the units' anticipated missions and levels of training readiness for each of the TOs. But the MDT2 Training Objectives Survey methodology could provide useful information on 'grass-roots' perception of the relative importance and 'trainability' of the TOs. In particular, ***the training value index (importance x effectiveness), rank-ordered across TOs provides a convenient supplementary source of information for making or revising decisions about TO priorities.***

The data base for the above could be drawn from a just completed training rotation and used by the unit commanders to plan future training. If derived from a cumulative archive of rotations, the data should probably be blocked by type of unit, e.g., reserve vs active to minimize variance across different types of units. However, the frequency and correlational analysis of the Training Objectives Survey data across years does not suggest the likelihood of a serious variance problem.

Identification of Training Problems. Illustrative training effectiveness data were used to show how the Training Objectives Survey methodology could help ***'red-flag' training objectives not adequately translated into scenario design***. The same methodology can be used to ***track improvements or decrements in training TO by TO***. This methodology can be generalized beyond MDT2 to any objectives-based, scenario driven training. In the illustration a cut-off value of 2.00 was used since this is the category value for 'minimal training requirements satisfied'.

Training System Assessment. The Training Objectives Survey data and analyses can be used in a number of generalizable ways to assess training systems. An overall ***frequency analysis can provide a temperature check on the relevance and effectiveness of the training*** provided by the system.

Summary measures of the importance and effectiveness indices can provide additional assessments. An example would be percent of TOs with average ratings > 2.00 (category value for minimal required training on the TO). The Training Objectives Survey data showed that in 1994 fourteen of seventeen TOs measured in the survey met the criterion for minimal training. In 1995 five out of twenty-five TOs assessed met the criterion. All the tasks in 1994 met or exceeded the criterion value of 2.00 for ***importance*** (category value for 'desirable training'. In 1995 the figures were twenty-two out of twenty-five.

Methodological Challenges for Distributed Training: Task Coverage Assessment.

Issue 1 'One size does not fit all': Initially we planned a single Training Objectives Survey instrument, but soon discovered the need to tailor survey forms to each site. Not all sites participated in all TOs. And even where TOs were common across sites, it became evident that each site might have a different perspective and/or involvement in the task.

This underscores the challenge of applying composite assessment indices to distributed training. By contrast, in a single-site, single Service environment, trainees and tasks-to-be learned are more homogenous. Therefore net training effective and value are more easily characterized.

Issue 2 Level of task description. Early in planning MDT2 exercises different perspectives emerged across the Services about the appropriate level of detail for describing TOs. We won't go into the minutia of those differences here. Suffice it to say that they pose a potential challenge in the DIS environment if it involves multiple organizations that have different experience with training task analysis. It's a problem that has to be worked out to accommodate the cooperating organization, while meeting the requirements of the training exercises.

Issue 3 Interpretation of responses. Negative responses on the ***importance*** or effectiveness scales will indicate a problem. But the nature of the problem will be unclear without comments either on the survey form itself or in subsequent interviews. The few comments we received in the Training Objectives Survey forms proved to be potentially very useful. They indicated specific defects in how TOs were implemented. The challenge then is to encourage comments or make provision for follow-up interviews. DIS aggravates the challenge because sites are geographically dispersed. The telephone conference is a possible solution, but needs to be arranged ahead of time.

Issue 4 Logistics and turnover. A problem encountered in MDT2 was that at Patuxent River and Armstrong, some players did not participate throughout the exercise week - for reasons beyond the control of the MDT2 staff. Therefore these players had only *partial* experience with the TOs relevant to their mission roles. This partial experience added a confounding source of variation to the assessment of the TOs. In a production version of MDT2 every effort should be made to avoid such turnover and its possible negative impact - on the training as well as the assessment. The problem is particularly aggravated by DIS, because players are at 'home' stations where it is easier to come and go than would be the case if all participants were brought together at a single site.

QUESTION 2A: HOW WELL WERE TRAINING OBJECTIVES COVERED BY MDT2?
Conclusions

1. Training objectives (TOs) were satisfactorily implemented by the MDT2 system and were important to include as training beyond that provided by the Services.

2. TOs selected by MDT2 CAS SMEs as key objectives were not judged by trainees to be more important (higher in priority) than the remaining training objectives.

3. The training value index (importance x effectiveness), rank-ordered across TOs provides a supplementary source of information for making or revising decisions about TO priorities and emphases.

4. The training effectiveness index can be used to identify TOs that are not being adequately implemented. It can also be used to track improvements from one training event to the next.

5. The frequency analysis, task effectiveness index, and importance index can be used to summarize training system effectiveness, relevance, and training value.

QUESTION 2B: HOW USEFUL WERE MDT2 INSTRUCTIONAL TOOLS AND METHODS?

Results

The results for Question 2B included taped interviews, transcribed into text. Interview comments were content analyzed using the format shown in Table 14. Comments were assigned to the categories of Table 14 as summary notes, direct quotations, or both. The intention was to capture each speaker's message as clearly and simply as possible, and to separate extended comments into focused categories. The analysis method was consistent with guidelines in Macey (1996). The results of this analysis are available upon request.

Table 14. Format For Content Analysis: Question 2B, Training Tools & Methods.

TRAINING DESIGN & MGT ISSUES	PARAPHRASED OR QUOTED COMMENTS						Summary
	Bn ST	CO TM	F16s	MULE	AFACs	O/Cs	
A. *FRONT-END PREPARATION*							
B. *TRANSLATION OF TOs;*							
1. Whole battle vs. battle slice							
2. Roles/structure: elements, echelons							
3.Training intervention							
4. Focus on key TOs							
C. *OTHER ISSUES*							
1. Collocation							
2. Variety of exercises							
3. Cross training							
AFTER ACTION REVIEW ISSUES	PARAPHRASED OR QUOTED COMMENTS						Summary
	Bn	CO TM	F16s	MULE	AFACs	O/Cs	
D. *AAR FEEDBACK TOOLS*							
1. 6X REPLAY							
2. STEALTH							
3. HARVARD GRAPHICS/UPAS							
4. TARGETS							
5. TOM							
E. *AAR DELIVERY METHODS*							
1. TELECONFERENCING							
2. FORMAT/ORGANIZATION							

QUESTION 2B: HOW USEFUL WERE MDT2 INSTRUCTIONAL TOOLS AND METHODS?
Discussion

A. *Front-End Preparation*

Participants commented on three aspects of MDT2 preparation: Training and orientation before exercise week, orientation on the first day of SIMEX week, and involvement in exercise planning or briefing on mission plans (e.g., commander's intent and scheme of maneuver). There was a cross-Service consensus that significant improvements were needed to prepare players and O/Cs prior to and during SIMEX week particularly on inter-Service topics. There were also Service-specific perspectives. These will be noted below.

Pre-SIMEX week. The Services agreed that they were sufficiently prepared to perform their own tasks, but felt that ***cross-Service coordination needed improvement***. The O/Cs at Ft. Knox, for example, felt that at least some key players - e.g., Task Force Commander and Air Liaison Officer - and all O/Cs would have benefited from visiting each of the sites. A number of participants across the Services suggested that more effort should be made in future applications of MDT2 ***to discuss the different Service roles, terminology, practices, and doctrine via telephone or video conferences before SIMEX week.***

The Air Force perspective varied somewhat from the above. ***Air Force participants indicated that it might have been helpful to 'brush-up' on CAS procedures since CAS was practiced infrequently.*** But they did not raise inter-Service preparation as a problem from their perspective. On the other hand, the Air Force O/C indicated he had helped write the scenarios and felt that was valuable inter-Service preparation for his role in the SIMEXs.

The read-ahead, including the Brigade Operations Order, ***drew criticism across the sites. It contained too much material and was not focused on the needs of specific sites.*** For example, the Air Force would have preferred a few pages that explained the general scenario, basics of the visual displays (what the icons represented), and SIMLIMs, especially those that constrained the pilots' tactics. ***MULE participants were confused about SIMLIMS vs. command or scenario limitations and they did not know how much initiative was expected.*** The company team players complained that they did not see the 'read ahead' except for the operations order. 'One size clearly did not fit all'. The implication is that more careful attention needs to be paid to the requirements of each site and 'read ahead' materials tailored to those requirements.

Players, especially at Ft. Knox, said they ***were unaware of TOs before SIMEX week.*** They would have benefited from a review of these at home station. The TOs were included in the 'read ahead' but apparently overlooked.

A strong consensus emerged on the inadequacy of the ***Brigade Operation Order (OPORD).*** It was described as ***not clear or complete, not well structured.***

> "A lot of the things we were scrounging around trying to figure out such as what is the Intel piece, what are the targets that have been nominated, what are the ACAs for this brigade level operation are not there in the order."

It was suggested that archived Bde Opords are available and these should be obtained in future MDT2 applications.

SIMEX week, day 1 orientation. A consensus emerged across the Services that ***Day 1 should be limited to orientation and preparation.*** An initial class of about 1 or 2 hours, walk-throughs (even MAPEXs) and opportunities for face to face discussions among participants, especially across the Services were suggested.

The initial orientation class should focus on SIMLIMS and Service specific terminology, procedures, and doctrine that the other Services needed to understand for the SIMEXs. A member of the Battalion Staff, for example, complained that he went through the entire week without understanding what an ATO is. That speaks to a shortfall in the preparatory training and not to a criticism of the staff member.

Briefing of Task Force mission plans to Non-Army participants. An Air Force F16 pilot felt that his participation would have been more effective if he had understood the Task Force commander's intent and scheme of maneuver, and the intelligence "picture." He recommended that the prebrief for pilots include such information. A second pilot thought this information had been deliberately withheld for instructional purposes. Such was not the case. ***Lesson learned: determine what information about task force mission planning should be made available to distributed sites, and how and when to provide that information.***

B. *Translation of Training Objectives*

The leap from training objectives (TOs) to the ***design of specific instructional experiences requires attending to the four major issues*** discussed below. The issues are not completely independent of each other. But they are separated somewhat arbitrarily to highlight key ideas.

1. Whole battle vs. slice of battle. What piece of a mission should be represented in an MDT2 environment? Not surprisingly, an underlying consensus across the Services and participants was that ***a clearly stated focus, a well defined purpose of training should determine what slice of battle to use in the training scenario.*** This idea is easier to state than to implement. Participant comments as well as observations during the exercises revealed a number places where the TOs and the goals of MDT2 did not match

exercise design. e.g., starting attack battles at the line of departure and long before the beginning of close air support. In subsequent exercises this was corrected. Missions were started closer to the CAS segment. But, in theory, this correction could have been made prior to exercise week on the basis of analysis rather than time-consuming experience during the week.

A solution to validating 'slice of battle' decisions is to systematically (e.g., through a check-list) verify those decisions against the TOs and general training goals. In fact this solution applies to each of the issues below. Tables which show the 'cross-walk' between TOs and training events (e.g., details of the exercise scenario) can facilitate this verification. Such tables are illustrated in Winsch, Garth, Ainslie, & Castleberry, (1996).

Participants - across Services and echelons - agreed that ***"slice-of-battle" decisions should be made within a building-block framework: "crawl", "walk", "run."*** In the crawl stage, for example, cross-Service comments suggested segmenting the MDT2 exercises even further than they were, e.g., separate "building-block" exercises involving the attack pilots and AFAC or MULE teams. This view is well supported by psychological research, (Teague & Park, 1996).

In any case, ***"slice-of-battle" decisions need to be constrained by available training time and resources***. In MDT2, Bde HQ was not played except for 'canned' Bde OPORDS, intelligence preparation of the battlefield was largely replaced by 'canned' information, and much of the 'normal' Battalion Task Force preparation was eliminated. The Fort Knox O/Cs emphasized that these decisions were made to conform to resources (as well as to provide greater focus on key TOs).

2. Roles and structure of unit elements and echelons. Once you've identified the "slice of battle", you need to define the structure of combat unit elements, the roles of specific players, and the consequences of reduced tactical realism (i.e., constraint on fidelity). ***Sub-dimensions of design include: (1) order of battle for friendlies and enemies, live vs. notional participation, "canned" vs free-play activities (especially for intel operations), and SIMLIM constraints on tactical behavior***.

The consensus on ***artificialities*** was that they ***are acceptable if they help focus training efficiently or if they're technically unavoidable, as long as players understand them ahead of time.*** A number of complaints about shortfalls in tactical realism were clearly related to misunderstanding or insufficient preliminary preparation and orientation of the players. The CAS pilots, for example, noted that AFAC participation was notably less than typical of other exercises such as Air Warrior. They were unaware that SIMLIMs at Patuxent severely limited the tactical play of the AFACs and their ability to support the F16 pilots. One AFAC also noted the under use of AFACs. He attributed this to inadequate

pre-exercise cross-Service coordination and lack of involvement in Task Force planning.

But a number of players suggested that ***more artificialities could have been used to advantage***. For example, the Bn Staff and CoTm players indicated that the CoTm players did not benefit from MDT2 training. They recommended ***replacing live CoTm players with a notional company team*** (i.e., company icon) and increasing the number of live scouts to compensate for visibility limitations.

Comments across the Services suggest some 'lessons learned'.

a. Document (with clear explanation) artificialities on unit structure or participant roles. Identify where tactical realism is constrained either because of a training strategy decision because of technical limitations.

b. Communicate these artificialities to participants well in advance of the training week, but at least during Day One orientation. Earlier communication is preferable since it allows for feedback and suggestions by participants. Also it will allow participants to alert the trainers and training managers to SIMLIMs that might otherwise be overlooked. Tannenbaum (1993) has shown that unrealistic expectations about what training can accomplish detract significantly from training value.

c. Be sensitive to the use of troops as 'training devices'. Consider replacing live troops with simulated participants/units where training benefit is expected to marginal.

3. Training interventions refer to changes introduced into 'normal' training scenarios to add challenge, stress, variety, uncertainty, or special training effects (Bjork, 1994). The exercise director may over-ride the 'normal' flow of events and activities by injecting an unexpected event or outcome to force rapid decision making. Interventions can also be used to exercise activities otherwise curtailed by the 'slice of battle' approach. For example, a sudden change in Air Coordination Areas (ACAs) was an effort to exercise decisions by the ALO and FSO which might have been required if MDT2 were run at the brigade or division level.

In MDT2 interventions were called 'trigger events'. Four 'triggers' were employed: breakdown in communication during a CAS engagement, change in ACAs 'imposed' by 'higher' command, increased artillery against friendly forces, and use of immediate CAS vs. planned CAS.

The ***reactions to 'trigger' events ranged from 'challenging and useful' to 'not needed'***. Furthermore the reactions varied with type of 'trigger' event. Change in ACAs was challenging to the ALO and FSO. It led them to consider three alternative solutions and then to make adjustments. The F16 pilots were neutral on heavy artillery and immediate CAS as special training interventions.

Communication breakdown drew the most consistently critical reactions. Participants were not clear whether the breakdown was part of tactical play or a failure in the simulation equipment. One player found this uncertainty distracting. In any case, there was some consensus that scenarios should not be complicated by 'triggers' early in the training week.

The interviews and research literature (Bjork, 1994), suggest that ***intervention design requires more systematic and detailed analysis than occurred in MDT2***. There the rationale for "intervening" was neither clearly defined nor documented. Furthermore, the trigger events, though derived from key TOs, were chosen arbitrarily and then not analyzed adequately to insure their usefulness and freedom from undesirable side effects.

The variety of interview reactions suggests that training intervention may be especially complicated in a distributed multi-Service environment. But they also suggest some ***rules of thumb for training intervention design.***

a. ***Design Training Intervention (TI) as part of the process of translating key TOs into scenarios***. Don't treat it as after-thought. For example, use TI as a mechanism for correcting or improving an initial translation. To do this go to step 2 below.

b. ***Identify and document the deficiency***. The deficiency may be a flaw in tactical realism or inefficiency in training strategy

c. ***Define and document explicit rationales for intervention*** and its design. At least three rationales come to mind.

(1). More closely match the operational environment by "introducing variability, delays, uncertainties, and other challenges the learner can be expected to face in the real-world job setting." (Bjork, 1994, P. 201).

(2). Introduce artificialities (i.e., distort tactical realism) to achieve special training effects or goals (Bjork, 1994, P202).

(3). Compensate for the artificiality of 'slice of battle' exercises by introducing effects from outside the 'slice' e.g., change in orders or priorities from higher headquarters.

e. ***Design the interventions***. Design should follow logically from the scenario deficiency and rationale for intervention.

f. Last but not least, ***assess the usefulness of the TI and check for undesirable side effects***. The MDT2 interview data uncovered some instances where interventions were viewed as having marginal value (e.g., immediate CAS for the F16 pilots and communications breakdown for a number of players).

4. Focus on key objectives. Based in part on Training Objectives Survey data from the 1994 exercises, seven TOs were designated as key training objectives for emphasis in the AARs. Interview dated generally supported the value of such focus. In fact the CoTm suggested highlighting key TOs in pre-exercise week preparations.

Furthermore, the players generally felt that the key TOs had been adequately represented in the MDT2 training system. There was some question, however, about how well Suppression of Enemy Air Defense (SEAD) was represented or even whether it should have been included since planning SEAD is primarily a Bde HQ function. Although SEAD is a very specific content area, questions about its inclusion and implementation raise a general issue about the selection of TOs for scenario design.

Interview comments suggest that perhaps SEAD should have been 'canned' or 'played' by O/Cs as were the Bde OPORD and the ATO. ***Training design implication: Take extra care to separate critical operational requirements which need to trained from those which can or should be 'canned'.***

C. *Other Training Design and Management Issues*

1. Collocation. ***Support for collocation was consistent across the Services***. The ***Mesa and MULE participants were especially enthusiastic about it. They felt that they could better understand each other's role in CAS because of the opportunities for face to face conversation before and after the exercises***. Since these units would not be together in combat, did ***collocation*** create tactical unreality and therefore adverse transfer? Probably not because the play of the mission itself was tactically realistic. The pre and post interactions (like AARs) constitute instructional leverage. A player at Patuxent River identified a side benefit: reduced turnover for the unit away from home station. Turnover was a severe problem at Patuxent River and confounded the learning curve.

The Knox O/Cs, in contrast, had reservations about collocation. They felt it defeated the purpose of DIS which was to avoid transportation of units. They anticipated future technology which could provide mobile and relatively inexpensive simulation facilities at any location where troops were home-based or stationed for duty. This latter vision, in fact was anticipated by a number of participants across the MDT2 sites.

An ***implication for future development*** is that ***ways should be sought to combine the face-to-face benefits of collocation with the cost benefits of distribution***, e.g., more informal contact, even by telephone, across sites, before exercise week. However, a more basic R&D effort would be to measure and explain the training value of ancillary, supplementary face-to-face contact during the training week.

2. Variety of Exercises. The CoTm players noted the lack of variety in mission design from exercise to exercise and the resulting predictability as exercise week progressed. The lack of variety was deliberately imposed to allow us to draw the learning curve. But the players' point is well taken for future training derived from MDT2 (Bjork, 1994. pp. 189-191). A lesson learned is: ***Vary the scenario design to minimize predictability from exercise to exercise,*** e.g., change the direction of attack and with it the mission objective.

3. Cross-Training. Participants at Knox noted cross-training as a coincidental, side-benefit of the MDT2 environment. CoTm players assumed multiple roles within the team and within the Bn staff. Bn staff also rotated roles. Comments indicated that these training opportunities were not normally feasible or available. ***Implication for training management: MDT2 may be a powerful environment for cross-training and developmental assignments to higher ranks.***

D. *After-Action Review Feedback Tools*

1. 6x Replay. ***Strong***, consistent ***support*** was expressed across echelons, Services, players, and O/Cs ***for the value of rapid replay of the Planned View Display (PVD).*** The PVD is a dynamic map of the battlefield which is used to recreate a training exercise. The players emphasized, ***however***, that ***rapid replay was useful only when focused on specific teaching points.*** Also, they recommended that better ways be found to highlight points made by the AAR leader, e.g., freeze-action and graphic highlighting of specific parts of the PVD. Their views are consistent with research support for the value of compressed time training techniques (Guckenberger, Guckenberger, and Stanney, 1995). Guckenberger et al specifically recommend the use of compressed time in distributed tactical training.

2. STEALTH. STEALTH is a display system that presents three-dimensional views of battlefield scenes that are simultaneously portrayed on the PVD. Players and O/Cs parted company on the value of STEALTH as an adjunct to the PVD during an AAR. ***Players perceived STEALTH to be distracting and ineffective.*** They indicated that it did not add to the information provided by the PVD. ***The O/C's, in contrast, opined that STEALTH was training-effective and added interest.*** Clearly, they were inaccurate on the second effect (added interest) since the players thought otherwise. And who can judge better than the players what they found interesting.

Research on multimedia training suggests that the O/Cs were probably inaccurate on the first effect (training-effective). ***Additional media are effective only if they supplement, complement, or reinforce the essential information provided by the primary training media and are contiguous in time (Etgen & Park, in press).*** The players' view was that the added media was a distracter. One O/C even suggested that the essential information

was not the same in the two media: "... the PVD gives you an architectural view, STEALTH gives you a dramatic view."

Training implication: Use of STEALTH as an adjunct to 6x replay should be avoided in AARs unless and until empirical or analytic evidence indicates otherwise.

3. Harvard Graphics/UPAS. The consensus across Services and echelons was that ***Harvard Graphics provided a useful tool for outlining AAR topics, summarizing teaching points, but was not useful for presenting battle damage assessment (BDA) data derived from the Unit Performance Assessment System (UPAS).*** The BDA data were considered to be neither accurate nor useful for correcting training deficiencies. Their connection to remediable performance problems was not clear, nor were attempts made to establish any connections. However, UPAS has been replaced by the Automated Training Analysis and Feedback System (ATAFS). ATAFS, an expert-based computer program supports the complete AAR process from description of events and outcomes to identifying what needs improvement and how to improve performance (Brown, Wilkinson, Nordyke, Hawkins, Robideaux, & Huyssoon; Meliza, 1995; U.S. Army Research Institute, 1996). ATAFS is designed for platoons and companies. However, similar systems exist for brigade and Corps exercises (Rankin & Gentner, 1996).

4. TARGETS. The Targeted Acceptable Responses to Generated Events or Tasks (TARGETS) instrument was one of two used by O/Cs to rate performance during the exercises (Dwyer et al.,1995; Fowlkes, Lane, Salas, Franz, & Oser, 1994). ***Reaction to TARGETS by O/Cs was varied but guardedly positive.*** The Ft Knox O/Cs found it manageable and useful for preparing AARs. The Air Force O/C was lukewarm. He found that TARGETS was good for documenting repetitive errors but did not reflect what he considered to be important assessment issues, e.g., "did tactics match the threat?". This comment suggests that check list measures during SIMEXs may need to be supplemented by post exercise assessments of performance against standards such as Commander's intent, enemy characteristics, and alternative tactical options. The AF O/C also found the work load heavy because of other roles. He suggested reducing the "laundry list" burden. The MULE O/C similarly felt that TARGETS "could have been easier." He did not use if to input to the AAR.

5. TOM. The Teamwork Observation Measure (TOM) was used by O/Cs to rate four dimensions of teamwork: Communication, Coordination, Adaptability, and Situational Awareness. ***O/Cs across the sites either ignored this form or found it not useful.*** The reaction is not surprising since TOM was constructed from the perspective of the research psychologist and based on psychological jargon. The underlying concept of going beyond go, no-go checklist measures may be valid. But TOM did not implement the concept successfully in MDT2.

E. *After-Action Review Delivery Methods*

1. Teleconferencing. Players and O/Cs split on this issue as they did on a number of others. The extremes were set by the **Bn** Task Force and the Fort Knox O/Cs. ***A consensus appeared across sites and participants on the value of telephone conferencing for the AARs. But the telecast of participants' faces was as strongly criticized by the Knox players as it was praised by the Knox O/Cs.*** Reactions at the other sites on this latter issue were split. The Air Force liked the video of faces, MULE participants did not.

Patuxent River participants did not comment. Research on multi-media effectiveness (Etgen & Park, in press) supports the negative views since there is no connection between information provided by facial features and vocalizations of participants - excluding the phenomenon of so-called 'body language.'

Training implication: Limit teleconferencing of AARs to voice transmission and video broadcasts of displays such as the PVD. If ATAFS or similar program were available that would be a suitable medium for distribution.

2. Format/Organization. The consensus across sites was that the organization of the AAR delivery was satisfactory by the end of SIMEX week and adequately balanced across the Services, i.e. sufficiently scenario driven. But it was too long and not focused enough on a few key teaching points. ***Participants recommended using the learning curve across exercises as a major organizing principle, in keeping with NTC practice.*** The Bn TF players were most critical in comparing the MDT2 AARs to those at the National Training Center. They commented on the professionalism, consistency, and focus of the NTC AARs. ***They recommended that future training based on MDT2 should have professional AAR cadres.***

Summary Of Comments On AAR Tools And Delivery

MDT2 participants agreed on the value of rapid replay of the PVD, the use of Harvard Graphics to outline AAR topics and teaching points, and telephone conferencing. They expressed negative views or strong disagreement on the value of battle damage or weapons effects data from the Unit Performance Assessment System (UPAS), use of Stealth to replay action, and use of VTC to show participants faces. Criticism of UPAS, however, has been overtaken by events. UPAS has been replaced by the Automated Training Analysis and Feedback System (Brown et al., 1995; U.S. Army Research Institute, 1996). Stealth and VTC were found, by the Fort Knox players, to be distracting. Table 14 outlines these and other conclusions.

Methodological Challenge for Distributed Training: Assessment of Instructional Tools and Methods.

Relativity is a problem here as it was for assessing training value. That is, effectiveness of a tool or method may be 'in the eye of the beholder', rather than an inherent quantity. Therefore, a primary challenge is to comprehensively tap variations in opinion and underlying reasons across sites and organizations. We found at least two sources of variation: one related to differences in site or Service perspective that could be traced to technical issues; the other was variation, between players and O/Cs, traceable to individual biases. An example of the first is the larger degree of criticism of AARs by Army players who were accustomed to very sophisticated reviews at NTC by professional cadres. An example of the second is the enthusiasm of the O/Cs, especially at Knox, for use of the stealth display during AARs compared to negative reactions by Army players.

The first source of variation provides useful information if the interviews or survey questions probe in sufficient detail to 'prescribe' improvements. Interview comments on the MDT2 AARs did provide specific ideas for improving the reviews. [See Discussion For Question 2b: Instructional Tools And Methods.] But our source of data was mainly open-ended questions. The responses across sites for specific training tools or features were very uneven: extensive at one site and perhaps negligible at another. This unevenness constrained conclusions about inter-Service value of the training tools and methods.

The structured survey provides a candidate solution to the challenge. Structured surveys are frequently used to focus on detailed characteristics of training media, methods, tools, or environments. They were not used here, but should be considered in future MDT2 or other DIS applications. A wide variety of formats and types of inquiries can be found in the survey literature; for example: Jarrett (1996), Johnson and Wightman (1995), McKeon (1994); Shlechter, Bessemer, Rowatt, and Nesselroade, 1994; Winsch, Atwood, Sawyer, Quinkert, Heiden, Smith, and Schwartz, (1994).

The second source of variation may be resolvable by invoking the research literature and widely accepted principles of training design. We did so in using the literature on multi-media research to support the Army players' negative views on use of stealth.

QUESTION 2B: HOW USEFUL WERE MDT2 INSTRUCTIONAL TOOLS AND METHODS

Conclusions

Major conclusions are presented below and detailed further in Table 15.

- Improved front-end preparation methods are critical to successful implementation of MDT2 to a production system

- Improved methodology for systematically translating TOs into scenario design is needed.

- Rapid replay, Harvard Graphics for topic outlining, and TARGETS are useful tools for AAR

- The value of STEALTH, TOM, and teleconferencing (particularly video displays of participants), as applied in this program is inconclusive at best.

Table 15. Conclusions About Training Tools and Methods.

TRAINING DESIGN & MGT ISSUES	CONCLUSIONS FROM CONTENT ANALYSIS
A. *FRONT-END PREPARATION*	Read ahead: reduce material; focused on site-specific needs. Improve Pre-SIMEX Week cross-Service coordination. Explain SIMLIMS during orientation. Spend Day 1 in orientation.
B. *TRANSLATION OF TOs:*	Need more systematic, accurate translation, site by site
1. WHOLE BATTLE VS SLICE OF BATTLE	Increase use of 'slice of battle' to support training focus within a building block strategy: "crawl, walk, run". Sharpen focus on multi-Service tasks. Constrain 'slice of battle' by available time and resources
2. ROLE/STRUCTURE OF ELEMENTS/ECHELONS	Replace live CO TM players with computer-generated icons. Add live scouts.
3. TRAINING INTERVENTION	Intervention design requires more systematic analysis and justification.
4. FOCUS ON KEY TOS	Strongly supported but need to begin focus before SIMEX WeeK. Need to separate critical operational tasks requiring training from those which can or should be 'canned'. Key TOs adequately represented in simulation, except for SEAD
OTHER DESIGN & MGT ISSUES	
1. COLLOCATION	Strongly supported, but ways need to be found to combine face to face benefits of collocation with cost benefits of distribution.
2. VARIETY OF EXERCISES	Vary the scenario design to minimize predictability from SIMEX to SIMEX.
3. CROSS-TRAINING	Benefit that emerged in MDT2-1995. MDT2 has strong potential for cross-training and developmental assignments to higher ranks.
AAR ISSUES	**CONCLUSIONS FROM CONTENT ANALYSIS**
D. *AAR FEEDBACK TOOLS*	
1. 6X REPLAY	Strong support when used to focus on specific teaching points.
2. STEALTH	Negative reactions among players, positive response from O/Cs. Negative reactions plus research do not support STEALTH as an adjunct to 6X replay.
3. HARVARD GRAPHICS(HG) /UPAS	HG was a useful tool for outlining AAR topics, summarizing teaching points, but was not useful for presenting battle damage assessment (BDA) data derived from the Unit Performance Assessment System (UPAS).
4. TARGETS	Useful for preparing AARs, but too long; missed 'important' assessments
5. TOM	Not useful for preparing AARs; ignored by some O/Cs
E. *AAR DELIVERY METHODS*	
1. TELECONFERENCING	Mixed reviews. Limit teleconferencing of AARs to voice transmission and video of displays such as the PVD.
2. FORMAT/ORGANIZATION	Need to follow NTC's AAR model: use learning curve across SIMEXs as an organizing principle and have professional AAR cadre.

LESSONS LEARNED ABOUT ASSESSMENT METHODOLOGY

Lessons learned are organized under assessment modeling, data collection general, and data collection -specific.

Development of an Assessment Model and Issues

We first considered using the widely accepted training assessment model of Kirkpatrick (1976, 1987), as extended by the Naval Air Warfare Center's Training Systems Division (Cannon-Bowers, Salas, Tannenbaum, & Mathieu, 1995; Kraiger, Ford, and Salas, 1993; Tannenbaum, Cannon-Bowers, Salas, and Mathieu, 1993). But we could not readily apply these models to MDT2 for reasons explained earlier in this report. The Kirkpatrick model was viewed as more appropriate for assessing classroom instruction. The NAWC/TSD model was viewed as more appropriate for basic research than for application since it is complicated and makes heavy data collection demands.

Lesson Learned: An alternative model, described at the beginning of this chapter is easy to communicate. Furthermore it yielded useful data. That model begins with two top-level questions. What is the value of MDT2 as a training system? How can it be improved? From these evolved assessment issues, measurement dimensions, and data collection instruments. More will be said about these shortly.

Data Collection - General

Comments added to survey responses

- Lessons Learned. In MDT2-94, survey respondents were encouraged to add comments to their responses. The comments which were made were useful. They helped make clearer the respondents' 'state of mind'. But they were few and far between. A review of Shlechter et al. (1994, Appendices A-C) suggested the addition of several blank lines after each opinion survey item. This addition, in MDT2-95, yielded an increase in the number of comments from 27 to 175.

- Recommendation: Encourage comments in Training Value Questionnaires by providing lined space after each survey item.

Linking items to dimensions.

- Lessons Learned. Because items were easy to generate, they were sometimes developed unsystematically. The linkages between items and assessment dimensions were less than explicit: items were related to dimensions only tangentially, or items related to multiple dimensions. Moreover, we relied solely on the judgments of project staff to define the dimensions, survey items, and linkages between the two. In retrospect, a pre-test sample from the population to be surveyed might have been used to help sort questionnaire items into value dimensions.

• Recommendation: Use a three-step approach to write survey items. "Brain-storm" survey dimensions and items as suggested by Babbitt and Nystrom(1989a). Prepare tables which show the linkages between items and dimensions. The tables should make clear items connected to multiple issues, as well as issues connected to multiple items. Finally, As part of the pre-test of items, use the pilot participants to verify linkage of items to dimensions.

Automation

• Lessons learned. Little or no use was made of automated procedures in data collection and processing. The results were inefficiency and extreme difficulty in keeping track of data. In MDT2-95 we considered creating machine scanable survey forms, but concluded that the labor and time required were not justified given the relatively small number of survey subjects. However, we did use a portable computer with a spread-sheet program on-site to record and summarize item response frequencies and to generate distribution graphs for the Training Value Questionnaire. This exercise was useful for securing data in electronic form and getting preliminary results.

• Recommendations: For a production training system, further, more systematic automation should be considered, particularly if surveys are to be used repeatedly for routine quality control. The ultimate system would include machine scoreable forms, on-site scanning capability, and entry into a computer data base for statistical analyses. For a one-time assessment with a small number of survey participants and items, machine scoring is not feasible because of the cost of designing and printing forms.

Pre-Testing

• Lesson Learned. MDT2 staff from each of the participating Services reviewed the survey forms. These reviews made clear the need for site-customized instruments and allowed us to revise specific items that were not adequately written. But logistic and resource limitations prevented us from pre-testing on a sample similar to the MDT2 participants. Pre-testing provides not only a change to improve items but also to validate their assignment to particular assessment model dimensions.

• Recommendation: Pre-test questionnaires, using subjects from the population to be surveyed. Instruct each respondent to apply the questionnaire to a familiar training system or environment. Then probe for assumptions, misunderstandings, and difficulties in answering each item. Finally, have the respondent sort the questionnaire items into the survey dimensions. Some care needs to be taken in defining and illustrating the dimensions.

Data Collection - Specific

Training Value Questionnaire. The purpose of the Training Value Questionnaire was to determine the participants' views on the value that MDT2 technology and methods might add to their existing training.

- Lessons Learned:

1. Respondents completed the Survey at the rate of two to four items per minute, double the estimated speed. Moreover, none of their comments or questions indicated difficulties in understanding or answering the items. Therefore, it seems feasible to use the questionnaire either for routine quality control in a production version of MDT2, or for ad hoc assessment of distributed training systems.

2. Survey items were not always sensitive enough to account for differences in perspectives that are peculiar to the DIS environment with its multiple sub-populations of players. For example, CoTm participants responded negatively to written questions about the value of MDT2 training. Their interview comments, however, indicated that they did see value for other echelons and units elements, but not for themselves.

- Recommendation. Where survey items are administered to subjects from different subpopulations,

Training Objectives Survey. The Training Objectives Survey assessed effectiveness of MDT2 for training 25 CAS task clusters (TOs). The questionnaire was adapted from an Air Force Study (Holstead, 1989). The clusters were developed to guide training design and evaluation of the MDT2 exercises. Each cluster (i.e., TOs) was a multi-Service collective action defined by a unique set of conditions, actions, and outcomes. Players and O/Cs rated each task on its importance for training and the effectiveness of MDT2 at providing that training.

- Lessons Learned:

1. Addition of a 'importance scale' to the original Air Force Training Objectives Survey was useful. With the training effectiveness scale, it yielded an index of training value.

2. Preliminary reviews of the survey form indicated that task clusters would prove difficult to assess (e.g., Control CAS Attack) because they were large, varied collection of actions. And, some of the participants experienced only some of the actions. Two steps were taken to cope with these problems.

a. Different sets of task clusters were used at five sites or echelons: (a) battalion staff (including TACP) and Knox O/Cs, (b) company commanders and platoon leaders, (c) TAC-A and Patuxent River O/Cs, (d) CAS pilots and Mesa O/Cs, and (e) laser

designator teams and NRad O/Cs. Respondents rated most of the tasks selected for them.

b. Participants were asked to check and rate only those parts of the task clusters which they had actually experienced. The resulting instruments appeared to be satisfactory.

3. The Training Objectives Survey may have been more difficult and time consuming than it should have been. Participants generally indicated in the interviews that they had not seen the MDT2 task clusters, i.e., training objectives (TOs)prior to SIMEX week. Nor were they even aware of them during the week and coming into the survey session. The survey required participants to read about a page of descriptive text per task cluster.

- Recommendations:

1. Customize Training Objectives Surveys to match the expected training experiences of the respondents at each DIS site.

2. Give participants the training objectives prior to SIMEX week. But command emphasis and active involvement of the training staff is need to insure that player understand the details, as well as the goals, of task performance.

Interviews. The interviews supplemented the written surveys in assessing satisfaction and reactions to training methods and media. They also solicited specific problems and benefits associated with MDT2. They were designed to capture comments that might shed light on written survey responses.

Utility of the interviews

- Lessons Learned. Interview data proved useful for helping to interpret the survey data. In addition, the MDT2-94 interview information helped us revise the assessment methods for Year 2. Comments were made in 1994 which suggested topics and probe questions for use in 1995. Interview data from 1995 was very useful in helping to explain some differences in Training Value Questionnaire Data between years and across sites. Site specific questions added in 1995 were especially helpful for this purpose.

- Recommendation: Use interviews routinely as part of a DIS assessment to supplement and highlight survey data. The interview is a potent tool particularly for assessing variations in feedback across sites and echelons.

Interview Logistics

- Lessons Learned:

1. The interviews were conducted in small, 'natural' groupings. For instance, O/Cs were interviewed separately from trainees. And trainees were separated by echelon or Service. The break-out by echelon and by trainee vs. O/C proved valuable. Group members stimulated comments among one another and expressed candid opinions freely. The Bn Staff was not a peer group since its ranks ranged from Sgt. to LTC. But, non commissioned officers did not hesitate to express opinions.

2. One hour appeared to be the minimal amount of time required. In several cases, interviews were cut short at the end of one hour because of other scheduled activities. The trainees and O/Cs seemed to appreciate the chance to talk about their experiences and reactions in a relaxed setting.

3. A variety of recording devices were used, ranging from hand held at Pautuxent River to a professional recorder with distributed microphones at Fort. Knox. The audibility of the recordings varied from minimally satisfactory to inaudible. Poor recording quality increased the complexity and cost of tape transcription.

- Recommendations:

1. Interview participants in small groups which are as homogenous as feasible. Allow one hour per group interview.

2. Every effort should be made to secure and use professional tape recording equipment. Clear and clearly audible tape recordings are critical to the effective processing of interview data. Notes are necessary but not sufficient. Make special efforts to secure quiet interview rooms and adequate recording technology. Last minute room arrangements, personal recording equipment, and throw-away portable recorders will not get the job done.

3. Furthermore, interviewers should be trained and checked out in use of the equipment. They should be trained in standard procedures for doing sound checks to insure the equipment is working and registering adequate sound volumes for each of the participants.

Interview Methodology

- Lessons Learned:

1. Interviewing methodology had a number of shortfalls. The tapes contained many inaudible or unintelligible responses. This problem reached its extreme in one case where a recording was blank for an entire session. Poor tape recording was aggravated by interview notes which were less detailed than anticipated. Sulzen (1995) describes similar problems in his Appendix B.

2. The notes were not sufficient for data processing purposes without follow-up discussions with the interviewers and access to recordings. The interviewers varied considerably in technique and interviewing skill. In the best cases, the interviewer actively intervened to clarify comments and summarize key ideas. In the worst cases, the interviewer did little beyond asking the major and the probe questions.

3. Several interviewers, particularly in Year 1 of MDT2, reported difficulty adhering to the wording and ordering of main and probe (i.e., follow-up) questions in the Interviewer Guide. They modified the interview protocol 'on the fly'. The result was a lack of uniformity across interview groups which made comparisons difficult. In MDT2, Year 2 more emphasis was placed on keeping to the interview script. As a result, comparisons across groups were easier to make.

4. In MDT2-94, attempts to have group members rank order good and bad features of MDT2 did not work. Trainees and O/Cs found it easier to highlight a few salient problems or good features. Interviewers did not persist in the rank ordering efforts.

- Recommendations:

1. Identify interviewers as early as possible and involve them in planning, developing, and revising interview methods and instruments.

2. Provide interviewers with detailed interviewing instructions, topics, and probe (i.e., follow-up) questions. To aid comparisons across groups of trainees and the accumulation of lessons learned, encourage interviewers to minimally follow the 'script'. But this should not preclude asking additional questions to clarify comments or encourage further information.

3. Provide training and rehearsal for interviewers in use of equipment and interviewing techniques. Interviewers especially need to be trained to spot responses that are inaudible, unintelligible, or irrelevant and to take corrective action. They should also be trained to participate actively by paraphrasing and summarizing key points. They should be encouraged to take notes of key points for use in summarizing and paraphrasing.

4. Avoid rank order questions. Instead ask the group to identify a few salient training problems or benefits where many ideas have surfaced for a particular question or phase of interview.

Interview Data Processing

- Lessons Learned:

1. Delays in reviewing tapes made the review and transcription more difficult because memories faded and interviewers at distant sites were more difficult to access for questions. This problem was aggravated because we did not pre-arrange for follow-up access to MDT2 players and O/Cs.

2. Content analysis was labor intensive and based on "penciled" notations in the tape transcripts. Electronic cut and paste was used to merge transcript passages with similar content classifications. More sophisticated content analysis would have been preferable.[14]

- Recommendations:

1. Put a high priority on data processing. Begin transcription as quickly as possible.

2. Arrange for follow-up access to survey and interview subjects, either through telephone calls or personal visits, if travel is feasible.

3. Semi-automate content analysis. Use a content-analysis computer program to conduct and document the analysis.[15] Alternatively transcribe interview data directly into a word-processing or data base table, assigning one transcript item per row. Classifications can then be assigned to each item. Finally items can be sorted by category. A data base provides more flexibility than a word processor table, but the table is easier to create and work with. It may be sufficient if simple category sorts are adequate.

Background/Experience in CAS. This survey instrument was used to collect military demographic information. It also asked whether the present job involves CAS, role in the MDT2 exercises, and CAS experience in academic programs and joint and Service-specific exercises. Forms for the Air Force personnel at Mesa, AZ, also asked for aircraft experience and highest current squadron qualification.

- Lesson Learned. Only one data gap surfaced. Respondents' military affiliations were not always clearly stated. Moreover, the same affiliations were sometimes expressed in different ways. Follow-up call to the various distributed sites were required to clarify affiliation entries.

- Recommendation: Obtain better demographic information by determining ahead of time the explicit titles of units and echelons that support the exercises. Tailor the forms to clearly indicate those organizations, either by providing instructions on standardized entries or a checklist of participating organizations.

RECOMMENDATIONS FOR USING MDT2 ASSESSMENT METHODOLOGY

Background

The purpose of this section of the report is to provide and illustrate recommended procedures for assessing user reactions to MDT2 as a training system. The illustrations come from the MDT2 trial demonstrations conducted in May of 1994 and February 1995.

The general purpose of assessment is to answer three questions. Does the training add value to the existing Service training 'pipelines'? (Does it meet the needs of the participating troops?). Is it working as it should? (What needs to be fixed?) And, finally, how useful are the component training methods and tools of MDT2?

Assessment supports total quality management (TQM), to which the Department of Defense (DOD) and the Services are committed. It assures that 'customer' expectations and needs continue to be met; that specific training design or management defects are uncovered; and that ideas for training system improvement are surfaced.

Why collect information on troop reactions? Earlier sections of this report explained how to collect process and product measures of performance and to examine changes in unit performance across MDT2 exercises. These changes in unit performance are the "bottom line". But participant opinion is equally critical.

1. Only satisfied customers will use simulation to train no matter how effective it is.

2. Therefore, user attitude towards training is an important consideration in cost-effectiveness analysis.

3. The customer knows best about product usefulness.

4. The customer has invaluable insights into needed product improvements.

5. Process performance data are subjective and outcome measures can be difficult to interpret. But combined with opinion information, performance data can help maintain high standards of training quality.

6. Satisfaction with training has a carry-over effect on post training performance.

Assessment Purpose

The purpose of assessment is to answer the following:

1. What value does MDT2 add to the Service training pipeline? This question is a "temperature" check, a check of "vital signs", a 'first-aid' diagnosis of MDT2 training system value and problems. It has five dimensions:

a. need for the training mdt2 can provide
b. credibility of the simulation
c. multi-Service value
d. role in training cycle
e. expected impact of mdt2 training.

2. How well does the system work? There are many ways to ask this question. A particularly efficient way is to ask how well were the training objectives for the exercises implemented? To what extent did MDT2 satisfy those objectives? Critical or negative answers can point to limitations in simulator capability (SIMLIMS) and deficiencies in scenario design.

3. How useful are MDT2 training tools and methods? MDT2 places special emphasis on training tools and methods to support front-end preparation, exercise conduct, and after-action review for 'long-haul', multi-Service training. The usefulness of these methods needs to be monitored so that they can be continually validated and improved or upgraded.

Users of and Uses for Assessment Data

Specific Units Accessing MDT2 Training. Bn CMDR and command staff can use training system assessment data as an additional tool for 'taking the unit pulse', communicating with subordinates about training needs. CMDR and staff can also use the data to determine future uses of MDT2 in unit's training cycle.

Cadres Managing the MDT2 training. A fully fielded version of MDT2 would be managed by a well trained and experienced cadre similar to the cadres at NTC or Nellis. They would be responsible for adjusting, revising, and updating MDT2 training methods and tools. They would also be responsible for identifying improved ways to translate training objectives into SIMEXS. Training system assessment data will provide a reference source of information for discharging these responsibilities.

Military Archival Clearing Houses. With the growing interest in networked, particularly 'long haul' simulation, it would serve the interests of the Services to collect and archive training system assessment data and lessons learned as they have for combat training centers. Two repositories for archiving, processing, and distributing the information are the Center for Army Lessons Learned (CALL) and National Simulation Center (NSC), both at Fort Leavenworth. Formal arrangements with these or other military clearing houses need to be explored.

Funding Agents. Finally, the military and political leadership will require continuing, documented evidence of the value of MDT2 in order to support it at the high levels that this type of training inherently requires to become institutionalized. Training system assessment data is appropriate to that purpose.

What Assessment Data to Collect and How

Overview. In the initial fielded applications of MDT2, assessment should focus on the needs of the specific units going through the training and on those of the managing cadres. In later applications assessment methodology would be modified to provide archival data and lessons learned for wider Service use. The guidelines provided below are geared towards initial applications of MDT2 fielded training.

The guidelines are intended as models, as templates, not prescriptions. Four procedures are recommended: collection of biographical information on each MDT2 player; application of a short, training-value questionnaire, administration of a questionnaire concerning how well training objectives were satisfied, and group debriefings to obtain more detailed information about what worked, what didn't work in the training.

Procedures for Training Value Assessment

Administer the Training-Value Questionnaire (Appendix A) across MDT2 sites, participating Services, and troops. The questionnaire should be given by a member of the MDT2 cadre after a minimum of two SIMEXs and related AARs. Schedule 2 minutes per item to allow time for written comments. For the 12 item questionnaire in Appendix A, schedule about 24 minutes.

How can you use the training-value data? What can you learn from it? Use the data to "take a unit's "pulse." For example, data from the 1995 demonstration with a an active Army unit, [reinforced by interviews] showed that company team members at Fort Knox saw a need for MDT2 CAS Multi-Service training at the battalion staff level. They felt their element and echelon did not benefit greatly. They made some excellent suggestions in the follow-up interview about how to design the scenario - next time around. In the 1994 demonstration, with a reserve unit at Fort Knox, we got a different "pulse reading". The company team found the training useful at their echelon as well as at the Bn staff level.

These "readings" can be used back at home station to stimulate discussions about training. The MDT2 cadre can use the training value results to help follow-on units plan the SIMEXs for their rotations. For example, they might recommend different training strategies for active and reserve component units.

Procedures for Assessing How Well Training Objectives were Satisfied

We anticipate that the commanders of each of the participating units and their staffs would select a subset of training objectives (TOs) from the a list of CAS TOs in preparation for their MDT2 rotation. With the assistance of the MDT2 cadre, the units would jointly develop SIMEX scenarios. Then as part of the assessment, the players, using the Training Objectives Survey would judge how well the TOs were satisfied by the simulation and the mission scenarios.

Administer the Training Objectives Survey across MDT2 sites, participating Services, and troops. Appendix A illustrates the 'critical' TOs that were used in the 1995 MDT2 Demonstration. The exact set of TOs for each site would be determined by the units participating in a rotation. The assessment sheets for these TOs would be assembled in customized packages as part of the preparation for the rotation. To improve the rotation's effectiveness and help the assessment, units should review, discuss, and map-exercise the TOs at home station.

Assign a member of the MDT2 cadre to give the questionnaire towards the end of the training week. We recommend Thursday afternoon (on a Monday to Friday schedule. Schedule 3 minutes per item. Each player will need to read a description of each TO, respond to check-list questions, and write comments. For the seven-item survey in Appendix A, schedule about twenty minutes.

How can you use this information and what can you learn from it? Use these data to check for agreement on the importance of the TOs for joint training and on how well MDT2 satisfied the TOs. Low ratings or disagreements are 'red flags'. Comments and debriefings may point to changes needed in simulation or scenario design on the next rotation.

Biographical Assessment

Hand out biographical data forms (Appendix A) for completion by all players. This form gets at the player's military background, but also his role(s) in MDT2. Since players may change roles across exercises, the biographical form should be given after the last exercise, or as late in the training week as possible.

How can you use this information and what can you learn from it? The primary use is to keep track of and organize the assessment data. It can also provide a record of the amount of cross-training which occurred during the rotation. If a player served in different roles, e.g., Bn CMDR as an O/C and S3 as Bn CMDR for one or more missions, those would be indicated on the background form. The unit commanders might use this information to document their use of cross-training. Such information could be useful in planning future training and in career development.

Group Debriefing Procedures

After the last SIMEX, players at each site should be debriefed in small groups. For example, in the MDT2 demonstrations, Bn staff were interviewed as one group, company team members were interviewed as a separate group. This breakout by echelon was designed to encourage candid discussion. If the company team were replaced by semi automated forces (SAFOR) (as suggested by assessment data from MDT2 demonstration in February 1995) different groupings would be appropriate, e.g., Bn staff in one group, scout units in another.

Appendix A contains guidelines for doing the group interviews. We recommend that these interviews do three things: check that significant training needs have been met and if not, why not; identify problems encountered; surface ideas to improve the training for the next rotation. Appendix A has sample questions about the three items above.

How can you use this information and what can you learn from it? Here are just a few ideas: (1) Use the comments to help guide follow-up discussions and planning back at home base; (2) compare comments across echelons and element, within Service. Look for serious disconnects. You may discover that not everyone is "reading off the same sheet of music."; (3) Do the same thing across Services, for the same reason; (4) Look for comments that may help explain negative responses on the Training Value survey and training objectives questionnaire. In the MDT2 Demonstrations we found the exit interview data to be a gold mine of information to help develop the MDT2 training and training management technology. Many excellent suggestions emerged.

REFERENCES

Babbitt, B. A., & Nystrom, C. O. (1989a, June). Questionnaire construction manual (ARI Research Product 89-20). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A212 365)

Babbitt, B. A., & Nystrom, C. O. (1989b, June). Questionnaire construction manual annex. Questionnaires: Literature survey and bibliography (ARI Research Product 89-21). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A213 255)

Bell, H. (1995). Symposium on distributed simulation for military training of teams/groups: The engineering of a training. Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting, San Diego, CA.

Boldovici, J. A., & Kolasinski, E. M. (1996, May). How to make decisions about device-based training: Elaborations on "what everybody knows". In Principal Scientist Colloquium (pp. 15-32) (ARI Special Report 26). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Bjork, B. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. P. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185-205). Cambridge, MA: MIT Press.

Brown, B., Wilkinson, S., Nordyke, J., Hawkins, R., Robideaux, B. & Huyssoon, S. (1996). Demonstrating the concept of an automated training analysis and feedback system (ARI Technical Report 1043). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD B215 579)

Burnside, B. L. (1990). Assessing the capabilities of training simulations: A method and Simulation Networking (SIMNET) application (ARI Research Report 1565). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A226 354)

Cannon-Bowers, J. S., Salas, E., Tannenbaum, S. I., & Mathieu, J. E. (1995). Toward theoretically based principles of training effectiveness: A model and initial empirical investigation. Military Psychology, 7, 141-164.

Center for Army Lessons Learned. (1995, February). Air Ground Operations School (AGOS). Combat Training Centers (CTCs) Bulletin No. 95-1. Fort Leavenworth, KS: Author.

Chen, T. T. (1995). A courseware instructional strategy model for equipment operation. CBT Solutions, November/December.

Converse, J. M., & Presser, S. (1986). Survey questions: Handcrafting the standardized questionnaire. Newbury Park, CA: Sage.

Crego, J. (1994). The exercising of critical decision makers and their teams: A simulation approach using digital video and a multi-station computer network. Proceedings of 16th Annual Conference and Exhibition on Interactive Systems for Training, Education, and Job Performance Improvements, 17-19 August, 1994, J.W. Marriott Hotel, Washington, DC.

Department of Defense. (1996, June). Recommendations for planning and conducting multi-Service tactical training with distributed interactive simulation technology (Four-Service Project Technical Report). Author: Washington, DC.

Dwyer, D. J., Oser, R. L., Fowlkes, J. E., & Meliza, L. (1995). Symposium on distributed simulation for military training of teams/groups: A case study of distributed training and training performance. Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting, San Diego, CA 1316-1320.

Etgen, M. P., & Park, O. C. (in press). Empirically-based principles for mutimedia presentation (ARI Technical Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Eyre, D. P. (1994). Cultural awareness and negotiation skills in peace operations. In D. R. Segal (Ed.), Peace operations: Workshop proceedings (ARI Research Report 1670). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A292 116)

Fowlkes, J. E., Lane, N. E., Salas, E., Franz, T., & Oser, R. (1994). Improving the measurement of team performance: The TARGETS methodology. Military Psychology, 6, 47-61.

Guckenberger, D., Guckenberger, L., & Stanney, K. (1995). Moving in time. Military Training and Simulation, Issue 2, 28-36.

Hamblin, A. C. (1974). Evaluation and control of training. London: McGraw-Hill.

Hawley, J. K., & Christ, R. E. (in press). Assessing the value added of new training technologies (ARI Research Report).

Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Hoffman, R. G., Graves, C. R., Koger, M. E., Flynn, M. R., & Sever, R. S. (1995, April). Developing the reserve component virtual training program: History and lessons learned (ARI Research Report 1675). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A296 153)

Holstead, J. R. (1989, June). Large scale simulation networking (SIMNET): Operational effectiveness appraisal (OEA), special project (TAC Project 89-190T). Eglin AFB, FL: U.S.A.F. Tactical Air Warfare Center.

Jarrett, P. A. (1996, February). Analysis of combat training center archive data for critical leader behaviors (ARI Technical Report 1038). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Johnson, D. M., & Wightman, D. C. (1995, November). Using virtual environments for terrain familiarization: Validation (ARI Research Report 1686). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A304 416)

Kirkpatrick, D. L. (1967). Evaluation of training. In R. L. Craig and L. R. Bittel (Eds.), Training and development handbook. New York: McGraw-Hill.

Kirkpatrick, D. L. (1976). Evaluation of training. In R. L. Craig (Ed.), Training and development handbook. New York: McGraw-Hill.

Kirkpatrick, D. L. (1987). Evaluation. In R. L. Craig (Ed.), Training and development handbook. New York: McGraw-Hill.

Knott, T. D. (1994). Planning & evaluating distance education: A collaborative approach. Memphis, TN: Diaphera.

Kosko, B. (1993). Fuzzy thinking: The new science of fuzzy logic. New York: Hyperion Press.

Kraiger, K., Ford, J. K., & Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. Journal of Applied Psychology, 78(2), 311-328.

Kraut, A. I. (Ed.). (1996). Organizational surveys, tools for assessment and change. San Francisco: Jossey-Bass.

Macey, W. H. (1996). Dealing with the data, collection, processing and analysis. In A. I. Kraut (Ed.), Organizational surveys, tools for assessment and change (pp. 204-232). San Francisco: Jossey-Bass.

McKeon, P. E. (1994). Evaluation of the airland battle management advanced technology demonstration prototype Version 1.2: Knowledge base assessment of the avenue of approach generation tool (ARI Research Product 94-07). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A280 241)

Meliza, L. L. (1993, November). Simulation networking/training requirements relational database (ARI Research Product 94-01). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A275 634)

Meliza, L. L. (1995, Winter). ATAFS: A first generation "smart" AAR system. ARI Newsletter, 5(4), 10-12. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Mirabella, A. (1995). Symposium on distributed simulation for military training of teams/groups: MDT2 system assessment and effectiveness. Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting, San Diego, CA, 1321-1325.

Moses, F. L. (1995). Symposium on distributed simulation for military training of teams/groups: The challenge of distributed training. Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting, San Diego, CA, 1306-1310.

Orlansky, J., Taylor, H. L., & Levine, D. B. (1996, in press). The cost and effectiveness of the Multi-Service Distributed Training Testbed (MDT2) for training close air support (IDA Paper P-3284). Alexandria, VA: The Institute for Defense Analysis.

Pinker, A., Samuel, A. H., & Batcher, R. (1995). Measures of effectiveness. Phalanx, 28(4), 8-12.

Rankin, W. J., & Gentner, F. C. (1996). A quick look back: After action review and debriefing methods. Modern Simulation and Training, 2, 30-36.

Rea, L. M., & Parker, R. A. (1992). Designing and conducting survey research. San Francisco: Jossey-Bass.

Schuman, H., & Presser, S. (1996). Questions and answers in attitude surveys, experiments on question form, wording, and context. Thousand Oaks, CA: Sage.

Shlechter, T. M., Bessemer, D. W., Rowatt, W. C., & Nesselroade, K. P. (1994, May). Evaluating the Unit Performance Assessment System's after action review displays (ARI Technical Report 997). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A281 712)

Stewart, J. E. (1994). Using the backward transfer paradigm to validate the AH-64 simulator training research advanced testbed for aviation (ARI Research Report 1666). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A285 758)

Sulzen, R. H. (1995, November). Training aids, devices, simulators, and simulations study (ARI Study Report 96-05). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Tannenbaum, S. I., Cannon-Bowers, J. A., Salas, E., & Mathieu, J. E. (1993, August). Factors that influence training effectiveness: A conceptual model and longitudinal analysis (ARI Technical Report 93-011). Naval Training Systems Center, Orlando, FL.

Teague, R. C., & Park, O. K. (1996). Environmental intensity, stress, and training (ARI Technical Report 1039). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A310 287)

U.S. Army Research Institute for the Behavioral and Social Sciences. (1996). The ARI "Products Book": Recent achievements in soldier-oriented research and development. Alexandria, VA: Author.

Winsch, B. J., Atwood, N. K., Sawyer, A. R., Quinkert, K., Heiden, C. K., Smith, P. G., & Schwartz, B. (1994, August). Innovative training concepts for use in distributed interactive simulation environments (ARI Research Product 94-16). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A285 585)

Winsch, B. J., Garth, T. H., Ainslie, F. M., & Castleberry, J. D. (1996, August). The development of structured simulation-based training for digital forces: Initial battalion staff-level efforts (ARI Research Product 96-08). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (AD A318 289)

APPENDIX A
Exercise Training Review Tools

This appendix describes the methods and tools to collect data about MDT2-CAS as a training system, in an After Training Review (ATR). The ATR provides feedback from the warfighter. The tools for obtaining that feedback make up the four sections of this appendix: (1) Biographical Data Form, (2) Training-Value Questionnaire, (3) Training Objectives Survey, and (4) Exit Debrief and Group Interview. As "experts" in their jobs, the O/Cs and trainees use the ATR to judge the success of MDT2-CAS. System developers need such judgments, combined with training performance data from the exercises, to gauge the system's quality and utility.

Procedures for After Training Review

- Hand out biographical data form for completion by all O/Cs and trainees at all sites.
- Administer the Training-Value Questionnaire to O/Cs and trainees at all sites
- Administer the Training-Objectives Survey to O/Cs and trainees at all sites.
- Perform Exit Debrief and Group Interview

Biographical Data Form

This form surveys each O/C's and trainee's military background and role in MDT2. Since trainees may change roles across exercises, the biographical form should be given after the last exercise or as late in the training week as possible. In addition, modifications to this form can address prior experience which may affect the O/C's or trainee's judgment.

Last four digits of SSN ____ Date __/__/__

BACKGROUND AND ROLE(S) MDT2

The purpose of this questionnaire is to determine your unit affiliation and the role or roles that you played in the past training week.

Demographic Information

1. Provide the following information about your present job:

   a. Rank/Grade: _____

   b. Job Title/Duty Position:

_____

_____

   c. Military Unit:

_____

_____

   d. [Check one]: Active Duty ____ Reserves ____ National Guard ____

   e. Does your present job involve planning, executing, or training CAS?

   ___ No.
   ___ Yes. Briefly describe how your job involves CAS:

_____

_____

_____

MDT2 Training

2. Indicate where you were physically located during the MDT2 training week:

   __ Fort Knox, KY __ Mesa, AZ __ Pax River, MD

3. What was your role in the simulated CAS missions? If more than one, assign "1" to the main role and then "2", "3" etc., to additional roles.

Army:
____ Bn/TF CMDR
____ S-3
____ S-2
____ FSO
____ Asst. S-3
____ FSE NCO
____ Co/Tm CMDR
____ Pl Ldr
____ FIST Chief ____ O/C
____Scout __Other:___

Air Force:
____ CAS Pilot
____ ALO
____ ETAC
____ O/C
____ Other: _____

Marines:
____ MULE operator
____ Frwrd Obsrvr
____ Ground FAC
____ TAC-A
____ O/C
____ Other:_____

Training-Value Questionnaire

This questionnaire asks about how well MDT2-CAS served a unit's need. It should be administered by a member of the MDT2 training staff after a minimum of two training SIMEXs and related AARs. Schedule two minutes per item to allow time for written comments.

Last four digits of SSN _____ **Date** _____

Instructions: Please give us your opinions about the value of MDT2 for muliti-Service training of close air support. For each of the following statements about MDT2, place an "X" in the space to indicate the extent to which you *agree or disagree*.

We would appreciate comments or examples, especially if you disagree with the survey item.

1. Need for Training Provided by MDT2

a. The opportunity provided by MDT2 to practice with personnel from other Services is necessary for training CAS.

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

b. MDT2 is a good training system for CAS because it focuses on critical training needs.

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

c. Given the opportunity, I would like to train with MDT2 on a periodic basis.

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

2. **Credibility**

a. **MDT2 can be an effective trainer for CAS with only a few, minor modifications.**

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

b. **A positive aspect of MDT2 is that it gives more realistic feedback to participants regarding CAS "kills" than is possible in field exercises or at Combat Training Centers.**

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

c. **I can apply more realistic CAS tactics in MDT2 than I can in field exercises or at Combat Training Centers.**

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

3. **Multi-Service Value**

a. **Experience on MDT2 made me better able to interact with members of other Services to plan for and execute CAS missions in combat.**

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

b. **Training with MDT2 gave me a better understanding of the jobs and roles of personnel from other Services in planning and conducting CAS.**

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

4. Role in Training Cycle

a. Experience on MDT2 better prepared me for field exercises on CAS missions, such as those at Air Warrior and NTC.

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

b. Training on MDT2 supplemented training in CAS that I receive within their military Service.

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

5. Expected Impact

a. The training that MDT2 provided can be applied directly to combat.

strongly agree	moderately agree	slightly agree	slightly disagree	moderately disagree	strongly disagree

Comments or examples? _____

b. Estimate the extent to which your experience with MDT2 has affected your ability to perform your role in a mission that uses CAS.

a.

no change in combat effectiveness	slight increase in combat effectiveness	moderate increase in combat effectiveness	large increase in combat effectiveness	extreme increase in combat effectiveness

b. Comments or examples? _____

Training-Objectives Survey

The O/Cs and trainees use this survey to judge how well the simulation and the mission scenarios for MDT2-CAS satisfied the training objectives. This survey should be given by a member of the MDT2-CAS training staff toward the end of SIMEX week. Schedule 3 minutes per item. Each player will need to read a description of each objective, respond to check-list questions, and write comments.

The set of all training objectives for MDT2-CAS appear in Table [1]. For routine training on CAS, commanders of each of the participating units and their staffs would select a subset of training objectives for emphasis. The assessment sheets for these objectives would be assembled in customized packages as part of training preparations. Use these data to check for agreement on the importance of the TOs for joint training and on how well MDT2 satisfied them. Low ratings or disagreements are 'red flags'. Comments and debriefings may point to changes needed in simulation or scenario design.

The inclosed set is an example from Fort Knox. Additional sets were tailored for the Air Force and Marine Corps participants. The inclosed set contains 7 key training objectives ("CAS tasks of special interest for MDT2") out of the total set of 25 shown in Table [1].

Bn Staff and ALO Survey Example

This questionnaire has two (2) parts. Part A contains CAS tasks of special interest for MDT2. Part B contains other CAS tasks. Most of the instructions are the same for both sets of tasks. Part A asks for one additional rating for the task and has more emphasis on comments than Part B.

PLEASE REMOVE THE FIRST TWO (2) PAGES TO USE FOR REFERENCE DURING YOUR RATINGS.

PART A

Purpose: The purpose of this survey is to determine the ability of MDT2 to provide training on selected close air support (CAS) tasks. These tasks conform to current CAS doctrine and emphasize the multi-Service nature of CAS. In addition to a short title, each task is described by conditions, actions, and outcomes.

Instructions: Read the complete description of each task and then answer the following questions:

1. **Participation or Observation**. Indicate whether or not you participated in or observed this task during the MDT2 demonstration. If not, skip the three remaining questions and go to the next task.

2. **Importance**. Rate the relative importance of receiving training on each task *in addition to training you have received* from your Service or in multi-Service military exercises. Use the following definitions to rate the importance of each task:

GRADE	Definition
A	Additional training for this task is essential.
B	Additional training for this task is highly desirable.
C	Additional training for this task is desirable.
D	Additional training for this task is somewhat desirable.
E	There is no need for additional training for this task.

3. **Ease of Use**. Please rate how easy it was to use MDT2 to train on the task. Place an "X" in the space provided to indicate the ease or difficulty of MDT2 use.

4. **Training effectiveness**. Rate the extent to which MDT2 provides effective training on the task. Use the following scale to rate the training effectiveness of MDT2 for each task:

GRADE	Definition
A	Provides all required training for this task.
B	Provides more than minimal essential, but not all required training for this task.
C	Provides minimal essential training for this task.
D	Provides some training, but less than minimum essential training for this task.
E	Provides none of the required training for this task.
F	Provides negative training for this task.
NR	Not rated or does not apply.

5. **Comments**. We would appreciate comments or examples, so that we can better evaluate MDT2. For example, if you checked A or B, did something stand out in the

training of the task being rated? If you checked D ("less than minimal training"), what is missing, or what would it take to increase your rating?

Part A contains 7 key tasks that directly pertain to battalion staff, including the ALO. Specific activities performed by the battalion staff are marked with a check mark (✓). Rate each using the definitions provided above. To remind you of these rating definitions, remove these instructions from the rest of the survey and keep them in front of you as reference as you rate each task.

Last four digits of SSN _____ Date _____

Instructions: Read the following description of task conditions, actions, and outcomes and then rate the task according to scales described in the instructions.

Title: Integrate CAS and other fire support elements with maneuver actions.

Conditions: The staff has received and understands the Bn/TF commander's concept of the operation. The staff has also been given the brigade priority/availability of fires.

Actions: The CAS plan is a subset of the fire support plan and conforms with the details of the Bn/TF maneuver plan and the Bde Fire Plan[b].

✓1. In consultation with the S-2, the S-3, and the ALO, the FSO generates a list of targets. In generating the list, staff must keep in mind that CAS targets must
   (a) have a purpose;
   (b) result from the intelligence preparation of the battlefield (IPB);
   (c) represent a high payoff target (HPT);
   (d) key on enemy, engagement areas, obstacles;
   (e) be based on the commander's intent and attack guidance; and
   (f) be manageable in number (i.e., 3-5 per company/team).

✓2. The staff prepares a Fire Execution Matrix, which is a graphic portrayal of fire support allocations. The matrix lists fire support elements by maneuver phases, thereby establishing execution responsibilities and coordination instructions.

✓3. The Bn/TF Commander approves the Fire Support Plan.

Behavioral Outcome or Product: The approved OPORD Fire Support Plan and the Fire Execution Matrix. Effects are continuous throughout the battle.

---

1. Did you participate in or observe this task being performed during the MDT2 demonstration?

   ____ No. Do not answer the questions on this page and proceed to the next task.
   ____ Yes. Respond to questions below.

2. Rate the Ta of receiving additional training by circling the corresponding letter:

   A B C D E

3. Rate how easy it was for your to use MDT2 for this task.

Very Easy	Easy	Borderline	Difficult	Very Difficult

4. Rate the training effectiveness of MDT2 for this task by circling the corresponding letter:

   A B C D E F NR

5. Comments or examples on your rating?

Instructions: Read the following description of task conditions, actions, and outcomes and then rate the task according to scales described in the instructions.

Title: Institute fire support control/coordination measures

Conditions: The Bn/TF has been provided with a coordinated fire line (CFL) and a fire support coordination line (FSCL). The TF commander has issued his concept of the operation including designation of high value targets (HVT) and high payoff targets (HPT).

Actions: In support of the maneuver plan, the staff develops a fire support plan that institutes the following fire support coordination measures:

✓1. designates restricted fire lines (RFL), restrictive fire areas (RFA), and no fire areas (NFA);

✓2. designates airspace coordination areas (ACA), critical friendly zones, and call-for-fire zones; and

✓3. establishes recognition and authentication procedures, a Fire Support Execution Matrix, and a Fire Support Attack Matrix.

Behavioral Outcome Or Product: The Fire Support Annex to the OPORD. Effects are continuous throughout the battle.

---

1. Did you participate in or observe this task being performed during the MDT2 demonstration?

   ____ No. Do not answer the questions on this page and proceed to the next task.
   ____ Yes. Respond to questions below.

2. Rate the importance of receiving additional training by circling the corresponding letter:

   A B C D E

3. Rate how easy it was for your to use MDT2 for this task.

Very Easy	Easy	Borderline	Difficult	Very Difficult

4. Rate the training effectiveness of MDT2 for this task by circling the corresponding letter:

   A B C D E F NR

5. Comments or examples on your rating?

Instructions: Read the following description of task conditions, actions, and outcomes and then rate the task according to scales described in the instructions.

Title: Initiate airspace coordination areas (ACA).

Conditions: CAS missions are planned (formal ACA) or CAS missions are to be executed on an immediate basis (informal ACA).

Actions: The FSO institutes measures to restrict fires into CAS airspace. To accomplish this, the following actions are taken:

1. If time permits, a formal ACA is deliberately planned. A formal ACA is a three-dimensional block of space in which aircraft are free to maneuver. Direct and indirect fires can be delivered, over, under, and around but not into the designated ACA. A formal ACA is the more desirable alternative because (a) it restricts less airspace; (b) allows tasking missions with the proper ordnance, sufficient time for planning, and integration of CAS mission with other missions; and (c) allows concentration of CAS, and avoids spreading CAS assets too thinly (piecemealing).

2. An informal ACA is the more likely alternative given the dynamic, fluid, and unpredictable nature of the battlefield. An informal ACA is simply a procedure for insuring separation of aircraft and surface fires. Fire support personnel should select a separation technique that requires the least coordination between air and firing units without adversely affecting the aircrew's ability to complete the mission safely. Aircraft and surface fire may be separated by distance or by time. Distance separation requires less detailed coordination than time separation but can be more restrictive for aircraft routing.

✓3. The FSO and ALO determine the appropriate airspace coordination measures. If the measures are too restrictive, the ability to achieve synergy of fire will be reduced. If the measures are too lax, the aircraft may be endangered by ground fire. Specially, (a) the ALO determines the characteristics of the aircraft/mission and translates that into airspace requirements and the time the ACA must remain in effect; and (b) the FSO considers the characteristics of the weapons available for indirect fires to avoid firing into the ACA.

✓4. The ACA measures are presented to the commander who determines their impact on his operational mission. The benefits of CAS may not be justified by the restrictions imposed by the ACA.

✓5. Given an approved ACA, the FSE plots and controls supporting fires in conjunction with the ACA.

6. Given an approved ACA, the TAC-A controls CAS aircraft to maneuver within ACA.

Behavioral Outcome or Product: ACA is documented in the Fire Support Plan Annex to the OPORD and/or the ACA plots in the FSE. The ultimate desired outcome is the least restrictive ACA in which CAS aircraft can operate safely and effectively. ACAs are also developed during the battle, particularly during immediate missions.

---

1. Did you participate in or observe this task being performed during the MDT2 demonstration?

   ____ No. Do not answer the questions on this page and proceed to the next task.
   ____ Yes. Respond to questions below.

2. Rate the importance of receiving additional training by circling the corresponding letter:

   A B C D E

3. Rate how easy it was for your to use MDT2 for this task.

Very Easy	Easy	Borderline	Difficult	Very Difficult

4. Rate the training effectiveness of MDT2 for this task by circling the corresponding letter:

   A B C D E F NR

5. Comments or examples on your rating?

Instructions: Read the following description of task conditions, actions, and outcomes and then rate the task according to scales described in the instructions.

Title: Incorporate SEAD in the fire plan.

Conditions: Enemy air defense assets have been identified and are located in the Bn/TF area.

Actions: To plan for the suppression of enemy air defense (SEAD):

✓1. FSO coordinates with the S-2 and S-3 to identify air defense assets and probable locations.

✓2. During planning, FSO plots probable locations for suppressive indirect fires. FSO also plans for subordinate commanders to use against air defense targets of opportunity. Indirect fires must conform with ACA.

✓3. During execution, FSO coordinates SEAD fires with the CAS delivery, such that fires impact one minute before the strike and continue for one minute after the aircraft have departed.

✓4. FSO synchronizes timing of air attack by relaying information about SEAD through the ALO to CAS pilots and the TAC-A.

Behavioral Outcome or Product: Enemy air defense capability is effectively suppressed. No aircraft are lost to enemy air defenses.

---

1. Did you participate in or observe this task being performed during the MDT2 demonstration?

   ____ No. Do not answer the questions on this page and proceed to the next task.
   ____ Yes. Respond to questions below.

2. Rate the importance of receiving additional training by circling the corresponding letter:

   A B C D E

3. Rate how easy it was for your to use MDT2 for this task.

Very Easy	Easy	Borderline	Difficult	Very Difficult

4. Rate the training effectiveness of MDT2 for this task by circling the corresponding letter:

   A B C D E F NR

5. Comments or examples on your rating?

Instructions: Read the following description of task conditions, actions, and outcomes and then rate the task according to scales described in the instructions.

Title: Update airborne pilots as necessary.

Conditions: Subordinate elements have reported changes in the tactical situation since the initial brief. Examples of tactical changes include movement of an enemy target, movement in the location of friendly forces, or a change in the status of enemy air defense artillery.

Actions:

✓1. The Bn/TF commander and his staff (including FSO) understand the input requirements for mission execution, and recognize changing tactical situations that impact on the mission.

✓2. The commander and/or staff normally direct the ALO to relay information to CAS pilots through the TAC-A. If close to air attack, this information may be communicated directly from the attack controller (ALO or ETAC) to the pilots.

Behavioral Outcome or Product: The pilots receive timely communication that accurately describes the changes in the tactical situation and its impact on the CAS mission.

---

1. Did you participate in or observe this task being performed during the MDT2 demonstration?

   ____ No. Do not answer the questions on this page and proceed to the next task.
   ____ Yes. Respond to questions below.

2. Rate the importance of receiving additional training by circling the corresponding letter:

   A B C D E

3. Rate how easy it was for your to use MDT2 for this task.

Very Easy	Easy	Borderline	Difficult	Very Difficult

4. Rate the training effectiveness of MDT2 for this task by circling the corresponding letter:

   A B C D E F NR

5. Comments or examples on your rating?

Instructions: Read the following description of task conditions, actions, and outcomes and then rate the task according to scales described in the instructions.

Title: Control CAS air attack.

Conditions: CAS missions has been approved and scheduled in to the Bn/TF sector. An forward air controller is airborne and on station.

Actions: In preparation for the attack, these actions are taken:

✓1. The ALO and TAC-A may confer to determine who will control the air attack. The final controller should be the one who is best able to control the air strike and observe its effects. This choice is likely to be the TAC-A, but the ALO or ETAC are both qualified air controllers who may be selected if they have a better line of sight.

✓2. The ALO identifies back-up elements (normally the FSO) in the event a qualified controller (i.e., the TAC-A, ALO, or ETAC) cannot control CAS aircraft. This step must be performed in case the communications between the aircraft and the Tactical Air Control Party (TACP) are disrupted.

✓3. The ALO insures (in conjunction with the FSO) that all primary and back-up elements have proper frequencies, call signs, correct CEOI, and are all operating in the secure/unsecured mode. Failure to do so may result in disrupted or ineffective communications.

✓4. The ALO insures that the TAC-A is on station and has communications with ground elements and the CAS aircraft. This step is necessary to determine whether or not the back-up controller must be used.

Behavioral Outcome or Product: ALO establishes communication with pilots and all elements involved in the CAS mission. The resulting communications between controller and pilot are accurate and timely.

---

1. Did you participate in or observe this task being performed during the MDT2 demonstration?

   ____ No. Do not answer the questions on this page and proceed to the next task.
   ____ Yes. Respond to questions below.

2. Rate the importance of receiving additional training by circling the corresponding letter:

   A B C D E

3. Rate how easy it was for your to use MDT2 for this task.

Very Easy	Easy	Borderline	Difficult	Very Difficult

4. Rate the training effectiveness of MDT2 for this task by circling the corresponding letter:

   A B C D E F NR

5. Comments or examples on your rating?

Instructions: Read the following description of task conditions, actions, and outcomes and then rate the task according to scales described in the instructions.

Title: Synchronize CAS attack with other direct and indirect fires.

Conditions: CAS aircraft are airborne and prepared to start mission.

Actions:

✓1. FSO directs time hacks to insure that suppression of enemy air defenses (SEAD) fire impacts 60 seconds prior to aircraft arrival in area of responsibility (AOR) and continues 60 seconds after their departure from AOR. Failure to do so could result in CAS aircraft receiving fire from enemy air defenses.

✓2. FSO directs time hacks to produce simultaneous effects of CAS with other fire support systems. Effects on enemy are greater when indirect fires are used simultaneously with CAS than when indirect fire and CAS are employed separately.

✓3. Through the FSO, Bn/TF commander synchronizes maneuver force attacks with the CAS attack. Simultaneous direct fire should increase the synergy of CAS and indirect fires even further.

Behavioral Outcome or Product: CAS aircraft are not hit by SEAD fire. Direct and indirect fires are coordinated with CAS attack to produce simultaneous effects.

---

1. Did you participate in or observe this task being performed during the MDT2 demonstration?

   ____ No. Do not answer the questions on this page and proceed to the next task.
   ____ Yes. Respond to questions below.

2. Rate the importance of receiving additional training by circling the corresponding letter:

   A B C D E

3. Rate how easy it was for your to use MDT2 for this task.

Very Easy	Easy	Borderline	Difficult	Very Difficult

4. Rate the training effectiveness of MDT2 for this task by circling the corresponding letter:

   A B C D E F NR

5. Comments or examples on your rating?

Exit Debrief and Group Interview

After the last SIMEX, MDT2-CAS staff should debrief O/Cs and trainees separately in small groups at each site. For example, in MDT2, Bn staff were interviewed as one group, company team members as a second group, and O/Cs as a third group. Such a break-out by echelon and type of participant encourages candid discussions. Allow one hour for the interview.

The following sections provide "Instructions for Conducting Interviews" and an "Interviewer Guide." The "Guide" contains broad open-ended questions as examples that were more detailed and site-specific in actual use.

Instructions for Conducting Interview

The following are general guidelines for conducting the post-demonstration interviews.

1. Conduct the interviews, after the final battle and AAR have been completed. If you cannot schedule an interview at this time, conduct the interview as late in the training as possible. For example, you may be able to schedule participant interviews before the final AAR. The interviewees should have finished all written surveys by this time, but if they haven't, give them time to complete the surveys before the interviews.

---

2. Tape record the interview. At the beginning of each recording, say the following information: (1) your name, (2) your location, (3) the date and time, and (4) the group you are interviewing (e.g., pilots, MULE O/Cs, etc.). Supplement the tape recordings with written notes in the space provided. If you make notes that are not on the interview forms themselves, please indicate on the notes the interview question to which they relate. Do a voice check on the recorder to be sure that the recorder is working and voices can be heard. For example, start the recorder, ask players to introduce themselves. In a very large group, ask a few people at the most distant locations to introduce themselves. Replay the recording. Make necessary adjustments If audio is poor, try some adjustments, like asking people to speak louder, move closer, or pass a microphone around.

3. Start the session with an introduction such as the following:

> "I would like to take some time for a discussion about the training that has taken place this week. The focus is on the strengths and weaknesses of this type of multi-Service training. I have a list of questions that I'd like to go through. After I read each question, let's discuss it."

Then read the first question.

4. Use the basic techniques of "good listening." That is, ask for clarification or examples, paraphrase your understanding of a respondent's comment. ***Do not*** let a comment go by if it is not clear, or if it contains acronyms. Ask the interviewee to repeat or paraphrase comments and explain acronyms. Keep in mind that clerical personnel with little or no subject matter expertise will be asked to transcribe the tape recordings.

5. Ask probe questions to elicit more detail or to get contrasting or corroborative opinions. The follow-up questions may also help to clarify special perspectives of participants with different views of the battles. [Suggested probes were provided to interviewers in the MDT2 assessment.] Use any follow-up questions that you think would provide useful information.

6. Maintain neutral appearance. Avoid interjecting your own opinion by word, gesture, or posture.

7. When the interview is over, put the date and your location on the tapes.

Interviewer Guide

Location _____ Interviewer _____
Group _____ Date _____

1. What significant training needs for your echelon and element were covered by the MDT2 training? How well were they covered? What needs were not covered?

2. What would you do to improve the following parts of the training system: (a) Preparation before coming here, (b) orientation on the first day of the rotation, (c) the simulation technology, and (d) the design of the mission scenarios for the SIMEXs?

3. What did you like most about this week's training? What did you like least?

APPENDIX B
Responses to Opinion Survey By Site and Item

Table B-1. Opinion Survey Across Sites-1994

ITEM	KNOX	MESA	DF-MULE*	PATUXENT
12/6	100% (22/22)	100% (4/4)	25% (1/4)	100% (1/1)
13/7	95% (21/22)	75% (3/4)	75% (3/4)	100% (1/1)
18/11	91% (20/22)	100% (4/4)	100% (4/4)	100% (1/1)
14/8	82% (18/22)	100% (4/4)	75% (3/4)	0% (0/1)
15/9	95% (21/22)	100% (4/4)	75% (3/4)	100% (1/1)
16/10	82% (18/22)	75% (3/4)	50% (2/4)	100% (1/1)
7/2	95% (21/22)	100% (4/4)	50% (2/4)	100% (1/1)
8/3	91% (20/22)	100% (4/4)	25% (1/4)	100% (1/1)
6/1	86% (19/22)	100% (3/3)	75% (3/4)	100% (1/1)
11/5	86% (19/22)	100% (4/4)	75% (3/4)	100% (1/1)
10/4	100% (22/22)	100% (4/4)	75% (3/4)	100% (1/1)
Totals	91% (221/242)	95% (41/43)	64% (28/44)	91% (10/11)

* Located at NPRDC, San Diego, CA _/_ = 94 vs. 95 numbering for items

Table B-2. Opinion Survey Across Sites-1995

ITEM	KNOX	MESA-AF	DF-MULE#	PATUXENT
12/6	83% (15/18)	100% (5/5)	100% (4/4)	100% (4/4)
13/7	72% (13/18)	100% (5/5)	75% (3/4)	50% (2/4)
18/11	69% (11/16)	100% (5/5)	100% (4/4)	100% (4/4)
14/8	56% (10/18)	80% (4/5)	50% (2/4)	25% (1/4)
15/9	63% (10/16)	100% (5/5)	100% (4/4)	50% (2/4)
16/10	53% (9/17)	80% (4/5)	50% (2/4)	25% (1/4)
7/2	83% (15/18)	100% (5/5)	100% (4/4)	100% (4/4)
8/3	86% (12/14)	100% (5/5)	75% (3/4)	100% (4/4)
6/1	94% (15/16)	100% (5/5)	100% (4/4)	50% (2/4)
11/5	65% (11/17)	100% (5/5)	100% (4/4)	75% (3/4)
10/4	100% (17/17)	100% (5/5)	100% (4/4)	100% (4/4)
Totals	75% (138/185)	96% (53/55)	86% (38/44)	70% (31/44)

#Collocated with Air Force at Mesa _/_ = 94 vs. 95 numbering for items

Table B-3. Opinion Survey at Knox by Echelon - 1995

ITEM	Knox	Knox Minus CoTm	Company Team
12/6	83% (15/18)	79% (11/14)	100% (4/4)
13/7	72% (13/18)	71% (10/14)	75% (3/4)
18/11	69% (11/16)	83% (10/12)	25% (1/4)
14/8	56% (10/18)	64% (9/14)	25% (1/4)
15/9	63% (10/16)	75% (9/12)	25% (1/4)
16/10	53% (9/17)	69% (9/13)	0% (0/4)
7/2	83% (15/18)	100% (14/14)	25% (1/4)
8/3	86% (12/14)	90% (9/10)	75 (3/4)
6/1	94% (15/16)	100% (12/12)	75% (3/4)
11/5	65% (11/17)	85% (11/13)	0% (0/4)
10/4	100% (17/17)	100% (13/13)	100% (4/4)
Totals	75% (138/185)	83% (117/141)	48% (21/44)

_/_ = 94 vs. 95 numbering for items

END NOTES

1 F16 pilots, for example, indicated in interviews that they had not received adequate information from Airborne Forward Air Controllers (AFACs) in OV10 simulators at Patuxtent River. This was related to simulator limitations at Patuxent River, which did not surface until after the test exercises had ended.

2 The 3/67 (3rd BN, 67 Armored Regiment)
2nd BDE/2nd Armored Division; 3rd Armored Corps

3 In May 1994 the MULE team was located at the Naval Research and Development Division (NRAD), San Diego, CA.

4 The MULE team was collocated with Air Force participants in 1995 to improve the efficiency of the MDT2 network. But, interviews indicated that collocation had many human factors benefits as well. For example, Air Force and Marine participants said that they could better understand each others' role in the exercises as a result of the collocation and opportunities to observe and have discussions with one another.

6 Babbitt and Nystrom (1989, Page 27).

> "It is sometimes useful to include one or more open-ended questions along with closed-end questions in order to obtain verbatim responses or comments that can be used to provide 'flavor' of responses in a report."

7 The expressions "task" and "training objective" were used interchangeably in the program and are so used here. Operationally the terms are defined by the descriptions which appear in the Training Objectives Survey.

8 Nominally, the total number of responses to the Training Objectives Survey was 31 (participants) x 17 (objectives) or 527 in 1994. The nominal total in 1995 was 31 x 25 or 775. However, in both years respondents had the option of passing on any objective they were not in a position to evaluate. Therefore, the total actual responses were less than 527 and 775 respectively.

9 Burside's methodology was also implemented by the Project Manager, Combined Arms Tactical Trainer, U.S. Army Simulation, Training and Instrumentation Command (STRICOM), as "Task Performance Support (TPS) Codes for CCTT.

[10] For the 1995 exercises, a panel of experts in multi-Service CAS operations identified seven key objectives for training emphasis. They selected tasks judged to be the most combat critical and requiring inter-Service coordination.

[11] For training objective Importance values of 4, 3, 2, 1, 0 were assigned to "essential", "highly desirable", "desirable", "somewhat desirable", and "no need". For Training Effectiveness,

values of 4, 3, 2, 1, 0, 0 were assigned to "all required", "more than minimal essential", "minimal essential", "some of the required training", "none of the required training", and "negative training" respectively.

[12] Note that a TO may have a very high combat "importance" but still have a low training "importance", e.g., the unit is already well prepared on that TO or the TO is better trained elsewhere.

[13] Differences in trainee populations across repeated applications of MDT2, however may confound use of the indices to assess training system improvements.

[14] For ideas on how to process verbatim comments see Kraut (1996, Pages 226-232).

[15] An example of such a program is FOCUS REPORTS, Professional Version 2.1 from Perseus Development Corp, 50 Aldrich Road, Needham, Massachusetts 02192,(617) 444-7355